DRAFT

Implementation and Assessment of a Multipurpose and Theory-Driven Emotion Awareness Tool.

With an Application to Distance Learning Settings.

PhD Thesis in Psychology
Université de Genève

By
Mattia A. Fritz
TECFA

Thesis supervisor
Prof. Mireille Bétrancourt
TECFA

Acknowledgements

I want to thank a few people.

Preface

This is an example of a thesis setup to use the reed thesis document class (for LaTeX) and the R bookdown package, in general.

Table of Contents

List of Tables

List of Figures

Abstract

The preface pretty much says it all.

Second paragraph of abstract starts here.

Dedication

You can have a dedication here if you wish.

Introduction

Imagine that you are co-authoring a text with a colleague for an upcoming assignment that you must submit in a course: you are at home, she can be anywhere in the world. You are using an online text processor, which allows you to edit the document simultaneously. This particular text editor has the feature of attributing animal names to authors, so that you are currently identified in the work-space as *anonymous hippo*, and your partner is *anonymous turtle*. The only thing you can see about the *turtle* is the flow of letters that stems from the position of her cursor in the document whenever she is typing on her keyboard, and *vice versa*. Imagine being in this situation: what kind of information about the *turtle* would you like to know to better assess how the learning task is going? And what kind of information about yourself would you like the *turtle* to know, for her to better assess how the learning task is going?

Now imagine that you are coding a project in an introductory programming course: you are facing lines of condensed and often mysterious syntax, whose execution sometimes results in the desired and celebrated outcome, but oftentimes leads to irking and obscure message errors. You are alone in front of your screen – riddled with software you barely know and dozens of tabs in your browsers with results to hazardous research queries – wondering whether your colleagues in the course are facing a similar fate. What information about them would help you better cope with this situation? What information about yourself would you like to share, which could be beneficial both to you and them?

You have certainly guessed from the title of the thesis that the answer put forward to these questions is: *emotions*. In the last few decades, emotions have been widely studied from different perspectives and with different methods, giving rise to the interdisciplinary field of affective sciences (e.g., Davidson, Scherer, & Goldsmith, 2003; Sander & Scherer, 2009). A consistent body of research has corroborated the importance that affective phenomena – such as emotions, feelings, motivations, moods, or preferences – play in various circumstances of life, both individually and collectively, and at different levels of processing (Adolphs & Anderson, 2018; Armony & Vuilleumier, 2013; Barrett, 2018; Colombetti, 2014; Damasio, 2018; A. H. Fischer & Manstead, 2016; Fontaine, Scherer, & Soriano, 2013; Pessoa, 2013; Rimé, 2005; Sander, 2013; Van Kleef, 2018). As a consequence of that, leading scholars in the field (Dukes et al., 2021) have recently considered legitimate to ask whether research has entered a new era: the era of *affectivism*.

> The conceptual, methodological, and technical advances made within the last few decades have demonstrated that affective processes are unquestionably enlightening when it comes to understanding both behaviour and cognition. While

> it will ultimately be the responsibility of historians of science to determine whether or not a new era has begun, given the undeniable impact of affective sciences on our models of brain, mind, and behaviour, it seems relevant to ask today whether we are now in the era of affectivism.
> — Dukes et al. (2021), p. 819

Even though *affectivism* may have had a slower start in learning and education sciences (Pekrun, 2005; Pekrun & Linnenbrink-Garcia, 2014a), there is nowadays a growing consensus in considering learning as the results of cognitive, social, and affective interactions (M. Baker, Andriessen, & Järvelä, 2013; Brackett, 2019; Immordino-Yang, 2016; Pekrun & Linnenbrink-Garcia, 2014a; Pekrun, Muis, Frenzel, & Götz, 2018). With respect to the role of affective phenomena more specifically, three intertwined research areas can be discerned. First, research investigating the effect of learning and teaching activities on learners' and teachers' affective experience (D'Mello, 2013; Pekrun, 2006; Pekrun & Linnenbrink-Garcia, 2014b). Second, research that, conversely, focuses on the effect of affective phenomena on learning processes and outcomes (D'Mello, Lehman, Pekrun, & Graesser, 2014; Hascher, 2010; Immordino-Yang, 2016). Third, research that investigates the existence of social and emotional competences that may be beneficial inside and outside learning contexts, and how to empower learners with those competences (Brackett, 2019; Brackett, Bailey, Hoffmann, & Simmons, 2019; Järvelä et al., 2016; Winne, 2015).

The role of affective phenomena in general, and emotions more specifically, is particularly challenging in computer-mediated learning environments – a term used throughout the text to encompass, at least partially, concepts such as advanced learning technologies (Graesser, 2013); computer-based learning environments (Moos & Azevedo, 2009); computer-supported collaborative learning (Dillenbourg, Järvelä, & Fischer, 2009); distance, remote or e-learning (A. W. Bates & Bates, 2005); and technology enhanced learning (Kirkwood & Price, 2014). The role of computational devices in both formal and informal education has become ubiquitous, especially in higher education, and a growing body of research is therefore devoted to investigate the interplay between affective phenomena and computer-mediated learning environments (Arguedas, Xhafa, Daradoumis, & Caballe, 2015; Cernea & Kerren, 2015; D'Mello & Graesser, 2015; Feidakis, 2016; Graesser, D'Mello, & Strain, 2014; Harley, Lajoie, Frasson, & Hall, 2017).

In this regard, part of the literature on the subject highlights shortcomings in computer-mediated learning environments due to the lack of para-verbal cues, usually available in non-mediated face-to-face interactions, which play a prominent role in various aspects of learning such as coordination, communication, sense of belonging or inter-subjective meaning-making (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Kreijns, Kirschner, & Jochems, 2003; Kreijns, Kirschner, & Vermeulen, 2013; Marchand & Gutierrez, 2012; Michinov & Michinov, 2008). In fact, even in co-located settings, the presence of a computer captures and directs learners' attention (Eligio, Ainsworth, & Crook, 2012), and therefore reduce the availability of cues, such as facial expressions or body postures, that are well known to convey emotional meaning (Bänziger, Grandjean, & Scherer, 2009; Martinez, Falvello, Aviezer, & Todorov, 2016; Scherer, Dieckmann, Unfried, Ellgring, & Mortillaro, 2019; Shuman, Clark-Polner, Meuleman, Sander, & Scherer, 2017). These shortcomings are further accentuated in distance learning, especially in asynchronous settings, where learners do not interact in real time, but have to cope with latency time and delayed exchanges (A. W. Bates & Bates, 2005; Jacquinot, 1993; Paquelin, 2011). There is a growing consensus in considering that such situations may benefit from some form of *social presence*, that is, the perception of others in the absence of *physical* presence (Gunawardena & Zittle, 1997; Jézégou, 2010; Lowenthal & Snelson, 2017). In this regard, several contributions highlight the importance of affective phenomena in providing a holistic experience of this social presence (Jézégou, 2010; Kirschner, Kreijns, Phielix, & Fransen, 2015; Lowenthal & Snelson, 2017; Rourke & Kanuka, 2009).

In the meantime, the literature also suggests that human beings' need for experiencing and expressing emotions transcends time and space (Derks, Fischer, & Bos, 2008; Parkinson, 2008; Rimé, 2005; Van Kleef, 2017). In this regard, the use of computers may provide alternatives, or even enhance emotional experience and communication, through dedicated or integrated artifacts (Boehner, DePaula, Dourish, & Sengers, 2007; Derks et al., 2008; Parkinson, 2008; Van Kleef, 2017). In other words, face-to-face interaction may not necessarily be considered as the golden standard to re-enact, for instance, by providing seamless audio/video connection (Buder, 2011; Janssen & Bodemer, 2013). On the contrary, emotion in computer-mediated learning environments may be conveyed through a wide panel of options, ranging from *simple* emoticons to autonomic recognition and representation through dedicated hardware and software (Arguedas et al., 2018; Baralou & McInnes, 2013; Cernea & Kerren, 2015; Glikson, Cheshin, & Van Kleef, 2018; Graesser et al., 2014; Harley et al., 2017). Computational devices open a wide spectrum of possibilities in the cycle of input, processing, and output of emotional information, including various means to gather data, persistence of data over time, visual-spatial displaying of data, as well as technologies that react or adapt based on learners' affective experience (Arguedas et al., 2018; Cernea, Weber, Kerren, & Ebert, 2014; D'Mello & Graesser, 2015; da Silva, de Souza Gomes, & de Almeida Neris, 2020; Fuentes et al., 2017; Grawemeyer et al., 2017; Kim, 2012; Leony, Muñoz-Merino, Pardo, & Kloos, 2013).

The choice of how to endow the computer-mediated learning environment with affect-related information and the reasons to do so depends on various criteria, with pedagogical and technical aspects often intertwined in mutual influence. In this regard, a growing body of research in the last decade has started to investigate the possibility to endow computer-mediated learning environments with emotional awareness, which is roughly defined as information about one's own emotions and/or about the emotions of other learners sharing the same computer-mediated learning environment (Arguedas et al., 2015; Avry, Molinari, Bétrancourt, & Chanel, 2020; Cernea & Kerren, 2015; D'Mello & Graesser, 2015; Eligio et al., 2012; Lavoué et al., 2020; Gaëlle Molinari, Chanel, Bétrancourt, Pun, & Bozelle, 2013). Contributions pertaining to this line of inquiry explicitly or implicitly share one ore more of the following underlying assumptions. First, it is posited that learners may benefit from being aware of their own emotions, that is, that their own emotions provide instrumental information for self-regulation of learning (Ez-zaouia, Tabard, & Lavoué, 2020; Lavoué et al., 2020; Gaëlle Molinari, Trannois, Tabard, & Lavoué, 2016; Winne, 2015). Second, it is assumed that emotion may convey instrumental information to others, for them to integrate that information into their own learning task, as a means to improve self-regulation, co-regulation and social regulation of learning (Eligio et al., 2012; Järvelä et al., 2015, 2016; Gaëlle Molinari, Chanel, et al., 2013; Winne, 2015). Third, it is also implied that emotion can be fruitfully conveyed in a computer-mediated learning environment, that is, the emotional information is made available and processed by users in the learning environment where learning activity and emotional awareness coexist in some form (Berset, 2018; Cernea & Kerren, 2015; Fuentes et al., 2017; Harley, 2015; Leony et al., 2013).

One way for learning activity and emotional awareness to coexist is through the use of an Emotion Awareness Tool (EAT). An EAT coalesces features of affect- or emotion-aware systems on the one side, and of awareness tools on the other. Affect- or emotion-aware systems (Calvo, D'Mello, Gratch, Kappas, & Graesser, 2015; Graesser et al., 2014; Grawemeyer et al., 2017; Harley et al., 2017) stem mainly from the interdisciplinary field of affective computing (Picard, 2000; Scherer, Bänziger, & Roesch, 2010) and are therefore driven, *in primis*, by computational modeling of affective phenomena (Marsella, Gratch, & Petta, 2010a; Scherer, 2010b). Awareness tools (Buder, 2011; Janssen & Bodemer, 2013; Janssen, Erkens, & Kirschner, 2011), on the other hand, emerged primarily from the field of computer-mediated collaboration and were later implemented in computer-mediated learning environments, especially through the interdisciplinary field of Computer-Supported Collaborative Learning (Dillenbourg et al., 2009; Kreijns et al., 2013; Miller & Hadwin, 2015; Suthers, 2006). Awareness tools have thus an intrinsic inter-personal stance, which emphasizes learners' need to build and update a mutual understanding of the situation at hand (Clark & Brennan, 1991; Roschelle & Teasley, 1995) as a means to maximize inter-subjective meaning-making, considered as a determinant of learning processes and outcomes (Dillenbourg et al., 2016; Janssen & Bodemer, 2013; Janssen et al., 2011; Suthers, 2006).

It follows that an EAT is a technological artifact, whose aim is to provide learners with emotional awareness in computer-mediated learning environments, under the assumption that its availability is beneficial to learners at the individual and/or collective levels (Arguedas, Daradoumis, & Xhafa, 2016; Avry, Chanel, Bétrancourt, & Molinari, 2020; Avry, Molinari, et al., 2020; Cernea & Kerren, 2015; Feidakis, Daradoumis, & Caballé, 2011; Lavoué et al., 2020; Gaëlle Molinari, Chanel, et al., 2013). Such a claim has received in the last decade an increasing interest in the affective, computer, and learning sciences from different theoretical, technical, methodological and empirical standpoints (*ibid.*). Applications of an EAT in computer-mediated learning environments may in fact range from providing learners with the possibility to self-report their emotions during a distance learning course as a means to increase self-reflection and self-regulation (Lavoué, Molinari, & Trannoi, 2017) to the implementation of a multi-modal system that provide emotional awareness through an unobtrusive brain-computer interface during computer-mediated collaboration (Makhkamova, Ziegler, & Werth, 2019). Despite the differences in application, though, research gravitating around an EAT usually aims at investigating which factors – intrinsic to the tool, deriving from the interaction between learners and the tool, between learners themselves, as well as between learners and the instructional design – determine whether and how emotional awareness may be beneficial in computer-mediated learning environments.

The thesis aims at extending this area of investigation by implementing and assessing a multipurpose, appraisal-driven emotion awareness tool based on self-reported emotions, which learners may express and perceive voluntarily and at any time while performing learning-oriented tasks in a computer-mediated environment. The implemented tool is also adopted – in an iterative interaction design and user-centered perspective (Cooper, Reinman, & Cronin, 2007; Lallemand & Gronier, 2017; Rogers, Sharp, & Preece, 2011) – in three empirical contributions, which are meant to investigate specific research questions, and at the same time inform the development of the EAT. In the reminder of this introduction, these intertwined objectives are first put into the general perspective adopted by the thesis. They are then specified individually in two sections that outline the methodological and empirical contributions respectively. The introduction ends with the outline of the manuscript, and with references to code and data, both of the implemented EAT and the computational manuscript of the whole thesis, which are made publicly available (Gilmore, Kennedy, & Adolph, 2018; Levenstein & Lyle, 2018).

Thesis Perspective

In the last few years, psychology has faced what has been widely referred to as the replication crisis (Anvari & Lakens, 2019; Chambers, 2017; Earp & Trafimow, 2015; Open Science Collaboration, 2015). What started as primarily a problem of statistical practices and inferences that undermined the reliability of experiments' results (Benjamin et al., 2018; Cassidy, Dimova, Giguère, Spence, & Stanley, 2019; Lakens et al., 2018; Nuijten, Hartgerink, Assen, Epskamp, & Wicherts, 2016; Open Science Collaboration, 2015) quickly turned in a wider methodological assessment of psychology's theories and overall practices (Chambers, 2017; John, Loewenstein, & Prelec, 2012; Makel, Hodges, Cook, & Plucker, 2019; McPhetres et al., 2021; Scheel, Tiokhin, Isager, & Lakens, 2020b). The term *credibility crisis* is sometimes adopted to refer to this wider problematic, for which, often, solutions are identified as pertaining to four intertwined categories (Maxwell, Delaney, & Kelley, 2017; Shadish, Cook, & Campbell, 2002; Vazire, Schiavone, & Bottesini, 2022). First, construct validity, which broadly refers to how concept are defined and measured (Scherer, 2005; Shuman & Scherer, 2014). Second, internal validity, which relates to identifying potential causal mechanisms explaining or predicting phenomena of interest (Grosz, Rohrer, & Thoemmes, 2020; Pearl, 2000; Pearl, Glymour, & Jewell, 2016; Pearl & Mackenzie, 2018a; Rohrer, 2018). Third, external validity, that is how knowledge built in the *micro-world* of experiments can be generalized to the *macro-world* of interest (Yarkoni, 2022). And, finally, statistical validity, which focuses on the link between domain knowledge and probability/statistical theory (McElreath, 2020; Rodgers, 2010).

Given that the field of emotional awareness in computer-mediated learning environments is relatively recent and still requires refinement on theoretical, methodological, and technical grounds (Cernea & Kerren, 2015; Eligio et al., 2012; Feidakis, 2016; Lavoué et al., 2020; Lavoué et al., 2017; Gaëlle Molinari, Chanel, et al., 2013), the thesis focuses primarily on construct and internal validity. The overall aim of the thesis can thus be divided in two intertwined objectives. On the one side, the methodological attempt to implement an EAT which maximizes the construct validity of emotional awareness. On the other side, the exploration of the internal validity of this same tool as an instrument that can be fruitfully implemented in empirical contributions investigating whether, how and why an EAT may be instrumental in computer-mediated learning environments. The two objectives are briefly outlined in the remainder of the section.

Methodological Contribution

Alongside other scholars (Arguedas et al., 2015; Brackett et al., 2019; Feidakis, 2016; Harley et al., 2017; Lavoué et al., 2020; Gaëlle Molinari et al., 2016), I reckon the way in which emotional awareness is implemented in a computer-mediated learning environment plays a prominent role in determining whether the EAT is adopted in the first place, and whether it serves its purpose in the wide spectrum of potential applications in which it may be deployed. Despite this variety, though, the role of an EAT may be broken down in two fundamental intertwined functions:

1. Whether, when, why, and how emotional information is *inserted/encoded* into the computer-mediated learning environments through the EAT, creating thus the necessary – but not sufficient – conditions for emotional awareness to emerge. In awareness tools, this function refers to *displaying* information (Buder, 2011; Schmidt, 2002) and is the equivalent of *expressing* emotion (Darwin, 1872; Scarantino, 2017; Van Kleef, 2017) when applied to an EAT.
2. Whether, when, why, and how emotional information is *retrieved/decoded* from the EAT and processed by learners, for emotional awareness to be enacted and hopefully contribute to learning processes and outcomes. In awareness tools, this relates to *monitoring* information (Buder, 2011; Schmidt, 2002), which corresponds to *perceiving* or *recognizing* emotion (Hall, Mast, & West, 2018; Hareli & Hess, 2010; Schlegel, Fontaine, & Scherer, 2017) in the context of an EAT.

One of the central claim of the thesis is that both functions may be better served with the interaction of three factors: (1) voluntary self-report of emotion; (2) implementing an emotion structure into the tool; and (3) moment-to-moment availability of emotional awareness in the computer-mediated learning environment. Each factor is outlined hereafter.

Voluntary Self-Report

I consider the opportunity that learners may best profit from emotional awareness through voluntary expression of their emotions (da Silva et al., 2020; Fuentes et al., 2017; Mortillaro & Mehu, 2015; Ritchie et al., 2016), as opposed to automatic emotion detection (Vallverdù, 2015). Self-report may encourage students to interpret, reflect on, extrapolate and construe meaning from emotion (Boehner et al., 2007; Fontaine et al., 2013; Lavoué et al., 2020) in a more flexible and adaptable process compared to the automatic detection of predefined *matches* from uni- or multi-modal sources (Calvo & D'Mello, 2010; Vallverdù, 2015).

Implementing an Emotion Structure into the Tool

I evaluate the possibility that an EAT may best serve its purpose if its functioning is driven by emotion theories (Fox, 2018; Moors, 2009; Sander, 2013; Scarantino, 2018; Scarantino & de Sousa, 2021; Scherer, 2022). In this regard, the thesis builds on appraisal theories of emotion (Moors, Ellsworth, Scherer, & Frijda, 2013; Roseman & Smith, 2001; Scherer, 2005, 2019a), and more specifically the Component Process Model theoretical framework (Scherer, 2005, 2009b, 2013b, 2019c). The central tenet of the model consists in considering that emotion is elicited and differentiated through an ongoing evaluation of events against a set of criteria, named sequential evaluation checks. This cognitive evaluation, or appraisal, then produces a synchronized, bounded physiological and behavioral pattern that may eventually coalesce into a conscious representation of the whole emotional experience, the subjective feeling, which can be used for intra-individual and inter-individual symbolic representation (*ibid.*). One of the assumption stemming from the theory is that, if one knows how the situation has been appraised by a person, it is possible to approximate what kind of emotional experience she will go through, and how she will name it Scherer & Meuleman (2013); Scherer & Fontaine (2018); Scherer & Fontaine (2019). This theoretical assumption is concretely implemented into the EAT through a parsimonious computational model, which, using an underlying affective space (Gillioz, Fontaine, Soriano, & Scherer, 2016; Scherer, Dan, & Flykt, 2006), bridges the cognitive evaluation of a situation and the resulting conscious emotional experience that learners may use for intra-personal and inter-personal emotion meaning-making (Grandjean, Sander, & Scherer, 2008; Scherer & Meuleman, 2013). Practically, the tool combines the dimensional and the discrete emotion approaches to emotion self-report (Mortillaro & Mehu, 2015; Scherer, 2005), which are dynamically linked by the probabilistic computational model. The result is an EAT that suggests the most likely subjective feelings to represent the emotional experience of a person, based on her rating on a set of appraisal dimensions (Fritz, 2015, 2016a; Fritz & Bétrancourt, 2017; Fritz, Bétrancourt, Molinari, & Pun, 2015). It is argued that implementing an emotion structure in this way may maximize emotion self-reflection and production of a strategic signal that can be conveyed to others (Scherer, 2007), even through a computer-mediated environment (Van Kleef, 2017).

Moment-to-Moment Availability of Emotional Awareness

Finally, I surmise that an EAT may increase its usefulness if learners are allowed to insert and retrieve emotional information at any moment (Graesser et al., 2014; Gaëlle Molinari, Chanel, et al., 2013). This refers to a distinction in instructional design between scripting and awareness tools (Miller & Hadwin, 2015). Within a scripting scenario, the learning activity is planned for learners to become aware of relevant information at specific and predefined moments (Dillenbourg, 2002; F. Fischer, Kollar, Stegmann, & Wecker, 2013). Conversely, awareness tools (Buder, 2011; Janssen & Bodemer, 2013) bestow learners with the responsibility of *producing and consuming* information at the very moment they see fit, either from an instrumental point of view (*I think this information is important right now*) or depending on the possibility to set their mind to it when coordinating multiple actions (*I* *can share or process this information now*). I argue that disposing of an EAT at any time in the computer-mediated learning environments may increase the chances of adopting it in the first place, but also to extrapolate meaning-making both at the intra-personal and inter-personal levels.

Methodological Objectives

The methodological objectives of the thesis are twofold. First, provide detailed information about the implementation of an EAT where these three factors coalesce. To meet this objective, I build on and extend previous work conducted in the Emotion Awareness Tool for Computer-Mediated Interactions (EATMINT) project (Avry, Chanel, et al., 2020; Avry, Molinari, et al., 2020; Cereghetti, Molinari, Chanel, Pun, & Bétrancourt, 2015; Chanel et al., 2016; Chanel, Molinari, Cereghetti, Pun, & Bétrancourt, 2013; Gaëlle Molinari, Chanel, et al., 2013). More specifically, a prototype of an EAT named Dynamic Emotion Wheel (Fritz et al., 2015), designed during an internship in the project and as the subject of my Master thesis (Fritz, 2015), will serve as the basis for the implementation of a functioning proof of concept.

Second, building upon the core computational model of the EAT, I propose a toolbox intended for scholars and practitioners alike, for them to dispose of a ready-to-use instrument, without the need for dedicated hardware and with minimal software setup (Cernea & Kerren, 2015). The toolbox allows the configuration of many aspects of the interface and inner functioning of the EAT, which may have theoretical, pedagogical and user-experience consequences. For instance, the toolbox provides the possibility to adopt different multi-dimensional affective spaces as the underlying structure of emotion (Fontaine, Scherer, Roesch, & Ellsworth, 2007; Gillioz et al., 2016; James A. Russell, 1980; Scherer et al., 2006). This has as a consequence the possibility to chose which appraisal dimensions users will be asked to rate, as well as the number and kind of suggestions they will receive from the system. The options in configuring the EAT are also intended to increase comparability of research within the affective science and education technology domains, as well as encouraging sharing of data and material in open formats (Lowndes et al., 2017; Nosek et al., 2015; Scheel, Tiokhin, Isager, & Lakens, 2020a; Scherer, 2005). Furthermore, the toolbox will also allow to chose between some graphical representation of emotion (Berset, 2018; Derick, Sedrakyan, Munoz-Merino, Delgado Kloos, & Verbert, 2017; Fritz, 2016b; Leony et al., 2013), which are meant to fulfill the perceiving-monitoring function of an awareness tool. This is the chance for researchers and scholars to decide how users will dispose of the emotional information inserted into the system, which, one again, may have important theoretical, pedagogical, and user-experience consequences.

To sum up, the methodological objectives of the thesis pertain to the implementation of a proof of concept, which aims at exploring whether an EAT with the aforementioned characteristics can be deployed and put at disposal of other scholars and practitioners interested in endowing computer-mediated environments with emotional awareness.

Empirical Contribution

As mentioned above, the empirical contribution of the thesis mainly aim at exploring the internal validity of the implemented EAT as an instrument to investigate intervening factors when emotional awareness is at play. Alongside other scholars (Avry, Molinari, et al., 2020; Cernea & Kerren, 2015; Eligio et al., 2012; Harley et al., 2017; Lavoué et al., 2020; Gaëlle Molinari et al., 2016; Ruiz et al., 2016), I reckon that computer-mediated learning environments can benefit from emotional awareness, for learners to dispose of meaningful information that they can harness for self-regulation, co-regulation, or socially shared regulation of learning (Janssen & Bodemer, 2013; Järvelä et al., 2015, 2016; Jézégou, 2010; Miller & Hadwin, 2015).

In this regard, the thesis proposes three empirical contribution related to the use of the implemented EAT. The three studies are briefly outlined hereafter.

Study 1

Considerable research has been devoted in the last few decades on how collaboration may best be sustained when learners share the temporal but not the spatial dimension (Dillenbourg et al., 2009; Järvelä et al., 2015; Kreijns et al., 2003; Roschelle & Teasley, 1995; Suthers, 2006). It is nowadays widely accepted that one of the determinants of the quality of interaction, as well as learning processes and outcomes, consists in learners building and maintaining a holistic representation of each others in order to maximize mutual understanding, coordination, regulation and inter-subjective meaning-making (Dillenbourg et al., 2016; Janssen & Bodemer, 2013; Järvelä et al., 2015; Gaëlle Molinari, Sangin, Dillenbourg, & Nüssli, 2009). More recently, some scholars have started to hypothesize that sharing emotions with the partner could, precisely, contribute to achieve this goal (Avry, Molinari, et al., 2020; Eligio et al., 2012; Feidakis, Caballé, Daradoumis, Jiménez, & Conesa, 2014; Gaëlle Molinari, Chanel, et al., 2013). The contributions investigating the matter have provided learners – often divided in dyads – with exactly the same emotional information: all the emotions shared by both learners, or none at all when using a control group without emotional awareness. When emotional awareness is provided, this ecological condition is consistent with the attempt to foster learners' mutual modeling, that is the act by which learners build and update a reciprocal holistic representation of what they are doing, thinking, or feeling (Dillenbourg et al., 2016; Engelmann, Dehler, Bodemer, & Buder, 2009; Gaëlle Molinari et al., 2009; Mirweis Sangin, Molinari, Nüssli, & Dillenbourg, 2011).

At a closer look, though, there are at least three intertwined – but still distinct – levels that may determine why a learner should consider to produce and peruse emotional information. First, the learner may be interested only in her own emotional experience. The fact of expressing how she feels may have benefit in terms of self-reflection and emotion self-regulation (Levenson, 1999; Torre & Lieberman, 2018). Second, the learner may be interested only in the emotions of the partners, and thus expressing her emotional experience is the reciprocal price to pay for everyone to dispose of this kind of information about others, which may be used as meaningful information for her own task (A. H. Fischer & Manstead, 2016; Rimé, 2009; Van Kleef, 2018). Third, the learner may be in fact interested in the *full* mutual-modeling experience, with, for instance, the possibility to compare her own emotional experience with that of the partners (Avry, 2021; Avry, Molinari, et al., 2020; Eligio et al., 2012; Gaëlle Molinari, Chanel, et al., 2013). The three situations incidentally coincide with the very definition of emotional awareness: information about one's own emotions and/or about the emotions of other learners. The first situation relates to the first part of the definition, whereas the second and third situation depends on the and/or distinction.

The first empirical contribution thus manipulate the interface of the EAT varying the use of and access to emotional information, under the assumption that this may change the concrete use of the EAT itself. In other words, learners will not produce and peruse emotional information at the same rate depending on what will be done of the emotion they express, and what emotional information they can access through the EAT. The three interfaces can in fact be placed on a scale which define how socially oriented they are, with the order of presentation reflecting the order of *social-orientation*. The underlying hypothesis states that the more socially-oriented is the interface, the greater use will be made of the EAT.

Study 2

As the recent events linked to the global pandemic have vividly purported, education cannot be transported *as-is* from the classroom to distance settings (TECFA, 2019). This certainly came as no surprise for scholars that have widely investigated potentials and limitations of distance learning in the last decades (A. W. Bates & Bates, 2005; Jacquinot, 1993; Jézégou, 2010; Paquelin, 2011; Sherry, 1995). A growing consensus has emerged about the need to attune the cognitive, pedagogical, and socio-affective dimensions in order to take advantage from distance learning and limit pitfalls, among which the sense of isolation resulting from the difficulty learners experience to connect with teachers or peers (Jézégou, 2010; Lowenthal & Snelson, 2017).

There is a growing consensus in considering that affective phenomena can play a prominent role in determining the overall attitude of students in distance learning settings (Feidakis et al., 2011; Jézégou, 2010; Lavoué et al., 2020; Lavoué et al., 2017; Lowenthal & Snelson, 2017). Even though emotion are nowadays considered by most scholars as short episodes elicited by specific events (Pekrun & Linnenbrink-Garcia, 2014b; Sander, 2013; Scherer, 2005), the cumulative emotional experience can have long lasting consequences on people's well-being and social interactions (A. H. Fischer & Manstead, 2016; Parkinson, Fischer, & Manstead, 2005; Rimé, 2005, 2009; Van Kleef, 2018). Scholars have therefore started to follow the emotional experience of learners during longer periods of time under the assumption that emotions may foster reflection and self-regulation, but also helps in projecting students socially and affectively in asynchronous settings (Jézégou, 2010; Lavoué et al., 2020; Lavoué et al., 2017; Lowenthal & Snelson, 2017; Ruiz et al., 2016)

In the same vein, the use of an EAT will be implemented in two classes of a blended Master in education technology at Geneva University. Students will be using the EAT during the distance periods in an introductory course about computational thinking and web development, which does not include in the instructional design collaborative projects

(Fritz & Schneider, 2019). The longitudinal plan aims at investigating whether the EAT is adopted in the first place, and whether the perception of its usefulness changes over time (Fitzmaurice, Laird, & Ware, 2011). The comparison between two different classes will also be the opportunity to explore whether the use of the EAT is mainly driven by individual attitudes, or whether some form of interaction may tight the individual use to the collective use (Dillenbourg et al., 2016; Janssen et al., 2011; Kirschner et al., 2015).

Study 3

The third empirical contribution takes advantage from the fact that the two previous ones have been conducted in very different settings: synchronous and collaborative the first, asynchronous and non-collaborative (or individual) the second. In the meantime, the toolbox was used to create instances of the tool who maintains a common *core*, which also happens to be the main characteristics of the EAT: self-report, computational model representing the emotion structure, and moment-to-moment availability.

The contribution thus combine data from both contributions and also from a previous usability test of the EAT performed in conditions very similar to the first study (Fritz, 2015). The aim is to conduct a sort of secondary data analysis (Weston, Ritchie, Rohrer, & Przybylski, 2019) or a small internal meta-analysis (Goh, Hall, & Rosenthal, 2016) to assess whether the characteristics of the EAT have been harnessed in neither, either, or both situations. The data-driven analysis will also be the chance to explore what type of information can be retrieved when data is collected through the same toolbox, but coming from instances of the EAT deployed in very different situations.

Empirical Objectives

The empirical objectives are one again twofold. First, respond to the specific research questions stated in each study, which may be of interest to scholars investigating whether and how endowing computer-mediated learning environments can be beneficial to learners. And second, collect first-hand data and know-how that can be directly exploited to assess the features of the EAT and, if possible, incrementally improve the overall toolbox.

Thesis Outline

The thesis is organized in four parts. Part I provide the overall theoretical background of the contribution. It starts from the general outlook of affective phenomena in learning and educational sciences and concludes with the specific inner mechanisms of emotion elicitation, differentiation, communication and regulation that may underlie the instrumentality of an emotional awareness tool in computer-mediated learning environments. One of the main aims of the first part is to sketch the intricate picture that emerges from the integration of articulated phenomena such as learning and affect in a multi- and inter-disciplinary field as education technology. Part I therefore aims at producing a reasonable overview on the subject at hand and justify to the best explanation the theoretical and methodological choices adopted for the proof of concept and the empirical parts of the thesis. More specifically, Chapter 1 proposes an overview of the inter-disciplinary convergence about the importance of implementing affective phenomena in learning as a means to improve both learning processes and outcomes, and of the different approaches through which this objective can be sustained. Chapter 2 discusses emotion awareness and emotion awareness tools in computer-mediated learning environments, proposing fundamental assumptions about the reasons why emotional awareness may be of use. The chapter also proposes an abstract model of emotional awareness, which put into perspective some of the numerous intervening factors that may determine whether and how learners take advantage from an EAT. The abstract model will be used as a reference throughout the thesis. Chapter 3 focuses on emotion theories and appraisal theories more specifically, under the assumption that an EAT may best serve its purpose if it implements an emotion structure, from which learners can take advantage in terms of emotion meaning-making and strategic signaling. The chapter will provide a definition of emotion that will inform the computational model at the core of the EAT.

Part II presents the implementation of the EAT and the toolbox built around it as a proof of concept, that is, a functioning device for which conceptual features have been prioritized over more fundamental practices in software development. Chapter 4 resumes previous work in the design of a prototype, named Dynamic Emotion Wheel, which will be extended both theoretically and technically in order to fulfill the multipurpose perspective. Chapter 5 illustrates the inner functioning of the parsimonious computational model that links appraisals criteria to the conscious emotional experience adopting a *k-nearest neighbor* approach, which can be scaled to multi-dimensional underlying affective spaces. Finally, Chapter 6 briefly outlines how a user-friendly toolbox has been built around this core model for scholars and practitioners to adapt the tool to their needs, as well as the features of the toolbox fostering comparability of studies and sharing of material and data. It is important to mention, though, that the proof of concept is presented before the empirical contribution for easing the general exposition of the thesis. As mentioned above, though, following an interaction design and user-centered process, Part II and Part III are highly intertwined: the technical and methodological implementation of the proof of concept has both informed and been informed by the empirical contributions. This means that Part II presents a *minimum viable product*, which is the result of the last iterative process, based on data and expertise gathered through the empirical contributions. At the same time, some of the technical and methodological elements resulting from the last iteration were not available for the empirical contributions. As a consequence, empirical contributions do not necessarily take full advantage or explore thoroughly all the available features. Furthermore, to keep this part of the contribution to a manageable length and accessible to non-technical readers, features of the device lacking a clear scientific impact will not be illustrated thoroughly. Interested readers may refer to the Code and Data section at the end of this Introduction for further details.

Part III comprises the three empirical contributions. Chapter 7 illustrates the experiment where the interface of the EAT is manipulated to create three different *flows* of emotional awareness. The experiment is characterized by a controlled environment, where participants were exposed to exactly the same stimuli through the simulation of a collaborative problem-solving task. The use of an eye-tracking device is also adopted to measure how participants interact with the different elements on screen when an EAT is integrated alongside a computer-mediated collaborative task. Chapter 8 reports the longitudinal study, which is performed in two different classes following a Master in education technology at Geneva University. The chapter introduces the Emotion Awareness Usefulness (EAU) scale, a tentative scale to measure emotional awareness instrumentality based on theoretical ground. An attempt is also made to investigate whether emotional competence, measured from a performance-based perspective, may be determined as an intervening factor in the use and perception of the EAT. Finally, Chapter 9 illustrates the data-driven analysis of the comprehensive assessment of the EAT. Different analysis will be carried out to assess the main features of the EAT, namely the presence of appraisal dimensions to evaluate the situation, as well as the use of an underlying affective space that informs the computational model in order to suggest the most likely subjective feeling to occur.

Part IV concludes the thesis with the general discussion of Chapter 10 and the Conclusion. The Appendices at the end of the manuscript only provide links to supplemental material which are available online.

Code and Data

The code and data for the toolbox as well as for the computational manuscript of the thesis – including thus data of the empirical contribution – are publicly available. Since the code span different repositories, but that those repositories may change the underlying technology or location over time, the different links are grouped in a single web page. The page is hosted on the TECFA server, active since the early 1990's, which has a long tradition of keeping URL stable over time. I hope this tradition will go on.

The link is the following:

- `https://tecfa.unige.ch/perso/mafritz/thesis/code-and-data/`

Part I

Theoretical Foundations

Chapter 1

Theoretical Convergence About the Instrumentality of Affective Phenomena in Computer-Mediated Learning Environments

This chapter provides an overview of the literature, combining different disciplines and perspectives, which corroborates the interest for endowing computer-mediated learning environments with affective phenomena. The chapter starts with an organized outlook of current trends in learning psychology and education sciences, which consider the interplay between affect and learning as a crucial interaction to foster learning processes and outcomes. Then, the chapter focuses on affect-aware systems and awareness tools respectively, for an Emotion Awareness Tool (EAT) shares similarities with both systems. The two systems, though, are driven mainly by different perspectives: affective computing in the first case, and Computer-Supported Collaborative Learning in the second. The emergence of an EAT with the characteristics stated in the introduction can therefore be compared and contrasted with the different perspectives provided in the literature.

1.1 Affect and Learning

The interplay between affect and learning is being investigated at various level of granularity, from the micro-level of neuroscience (*e.g.*, Immordino-Yang, 2016) to the macro-level of how education curricula and institution should reform their programs in order to take affective experiences into account (*e.g.*, Brackett, 2019). Rather than a layered analysis, though, this section proposes a selection of contributions – privileging research related to learning with technology whenever possible – organized in three categories, which are nevertheless not meant to be neither mutually exclusive nor exhaustive: (1) research that focuses *primarily* on the effects of learning processes and outcomes on affect; (2) research that focuses *primarily* on the effects of affect on learning processes and outcomes; and (3) research that focuses *primarily* on how either or both directional effects can be mediated or moderated by various forms of socio-affective competences. The main aim of the section is to depict a growing cross- and inter-disciplinary interest in considering learning as the result of cognitive, social, and affective interactions. This will provide the occasion to pinpoint how emotional awareness conveyed through an EAT relates to these research areas.

1.1.1 The Effects of Learning on Affect

Research investigating the effect of learning on affective phenomena is primarily concerned in determining what affective reactions are elicited by learning processes and situations, and what are the reasons and mechanisms that concur in eliciting such reactions. According to Pekrun (2005, 2006), for instance, the scientific study of emotion in education emerged in the last few decades by extension of previous work that was primarily focused on reducing anxiety during test exams. In an effort to broaden the perspective on the interplay between emotion and learning, Pekrun set forth an overarching theory, named the *Control-Value Theory of Achievement Emotions* (Pekrun, 2006), which accounts for eliciting mechanisms and consequences of learning situations where some form of achievement is at stake (*e.g.*, passing an exam). The theory states that learners evaluate their learning experience according to two criteria: the subjective control that learners feel over the learning process, and the intrinsic value that they confer to the achievement outcome. Different emotions, such as *frustration*, *enjoyment*, or *boredom*, are then elicited based on the specific evaluation that learners make about the situation. The evaluation may be retrospective, perspective, or ongoing. The Control-Value theory is depicted more thoroughly in Chapter 3, but is presented here as a line of inquiry that aims at explaining why and how certain emotions are elicited by learning processes. In this regard, it is also worth mentioning that achievement emotions are only one possible *kind* of emotion that may be elicited in learning. Pekrun and Linnenbrink-Garcia (2014b) proposes a taxonomy articulated around four *families* of emotions related to learning and educational settings:

1. *Achievement emotions*, already mentioned, elicited by activities or outcomes that refer to standards of evaluation. Examples of achievement emotions are *pride* or *shame* that may occur depending on the result of an exam.
2. *Epistemic emotions*, elicited by processes that learners enact to assimilate and accommodate new information in the intent to build knowledge. Examples of epistemic emotions are *curiosity* or *surprise* emerging from incoming information that may contradict pre-existing knowledge;
3. *Topic emotions*, elicited directly by the content of the learning activity, independently of achievement or epistemic processes. Examples of topic emotions are *empathy* towards the fate of a character in a novel, or strong affective reactions that may arise from historical or political events;
4. *Social emotions*, elicited by inter-personal dynamics that are pervasive in learning and educational contexts, such as the relationship with colleagues or teachers. Examples of social emotions are *admiration* for a colleague, who was of help in solving a problem, or *anger* towards a free-rider in a group assignment.

One important element to retain from this taxonomy is that the same *discrete* emotion may pertain to different *families* according to what is the object focus of the eliciting situation (Pekrun, 2006; Pekrun & Linnenbrink-Garcia, 2014b). For instance, *frustration* may be triggered by the result of an exam (achievement), by the self-oriented perception of inability to understand information (epistemic), by the unwillingness to manipulate scientific equipment in a science class (topic), or by the difficulty in explaining to others one's opinion (social).

Another research approach pertaining to the effects of learning on affective phenomena consists in focusing on the kind and frequency of specific affective phenomena emerging from learning experiences, without necessarily proposing an eliciting mechanism. D'Mello (2013), for instance, conducted a selective meta-analysis on affective states experienced during individual learning with technology. Combining 24 studies from 5 different countries and from different level of formal education, the author computed the relative frequency of different affective states. Measures were recorded from different sources, from self-report to retrospective video analysis by trained judges. Results show that only one affective state, defined as *engagement/flow*, was consistently frequent in most of the studies. More traditional affective reactions such as *anger*, *anxiety*, *boredom*, *fear*, *frustration*, *happiness*, *sadness* and *surprise* were either infrequent, or subject to great heterogeneity between studies.

Reis and colleagues (2018), on the other hand, performed a similar meta-analysis, but encompassing affective states during Computer-Supported Collaborative Learning (CSCL) (Dillenbourg et al., 2009; Stahl, Koschmann, & Suthers, 2006; Suthers, 2006). Contrary to D'Mello's contribution (*ibid*), this second meta-analysis integrates also a collective, human-to-human interaction perspective. Starting from an initial corpus of over one thousand contributions between published articles and conference papers, the authors extracted 58 papers. From these, they then computed the number of papers citing different discrete affective states, belonging to different categories of affective phenomena: personality traits, emotions, moods, and socio-emotional factors. With respect to affective states identified as emotion, the authors tallied 29 discrete emotions that appeared in a minimum of two and a maximum of eight papers. The most frequently cited discrete emotions are: *angry*, *confused*, *disgust*, *pride*, and *tired* (in four papers); *curious*, *excited*, *interested*, *joy*, *sadness* and *surprise* (in five papers); *anxious*, *bored*, *relaxed* and *anger* (in six papers); *fear* (in seven papers); and *frustration* (in eight papers).

This brief selection of contributions corroborates the fact that learning does indeed elicit affective phenomena, but that those phenomena varies greatly depending on a number of situated factors. Research is therefore implicated in determining the mechanisms of elicitation, as well as explaining the potential causes of variance in affective experiences according to different learning settings. In this regard, an EAT may be deployed in different learning contexts, for instance in individual vs. collaborative activities.

1.1.2 The Effects of Affect on Learning

Another line of inquiry, which is highly intertwined with the one illustrated in the previous subsection, aims at investigating whether affect in general, and specific affective states more specifically, have effects on learning. One common line of inquiry concerns the identification of emotions that may foster or hinder learning processes or outcomes (D'Mello et al., 2014; Graesser et al., 2014; Harley et al., 2017; Pekrun, Goetz, Titz, & Perry, 2002; Ruiz et al., 2016). Research in this area may vary greatly depending on the level of analysis and the affective phenomena they target.

With respect to the level of analysis, it may concern the interplay between affective and cognitive processes such as attention, perception, memory and decision making, which are deeply implicated in learning processes and outcomes (Brosch, Scherer, Grandjean, & Sander, 2013; Immordino-Yang, 2016; Pessoa, 2013). Research may target higher-order phenomena, such as creativity or problem-solving (Avry, Chanel, et al., 2020; Avry, Molinari, et al., 2020; M. Davis, 2009). Yet another target is constituted by inter-personal processes, in which affective phenomena may regulate or undermine collaborative situations (Andriessen, Baker, & der Puil, 2011; M. Baker et al., 2013; Järvenoja, Volet, & Järvelä, 2013).

Concerning the kind and nature of affective phenomena, research often adopts a dichotomy between *positive* and *negative* affective phenomena (Harley et al., 2017; Hascher, 2010; Pekrun et al., 2002; Rowe & Fitness, 2018; Ruiz et al., 2016; Shiota, Sauter, & Desmet, 2021). For instance, Ruiz et al. (2016) propose the Twelve Emotions in Academia Model (TEAM), which is meant to foster individual introspection of emotion by providing learners with a self-report tool. The tool comprises 12 discrete emotions equally divided in *positive* and *negative* emotions. The *positive* emotions are *enjoyment*, *hope*, *pride*, *confidence*, *excitement*, and *interest*. The *negative* emotions are *anxiety*, *anger*, *shame*, *hopelessness*, *boredom*, and *frustration*.

This sort of dichotomous categorization, even though intuitive at first, is nevertheless problematic at a closer look. Shiota, Sauter and Desmet (2021), for instance, identify three different ways in which the polarity of an affective phenomena is conceptualized in the literature:

1. the *valence* of subjective feelings, which usually refer to an assessment of the situation as pleasant or unpleasant, even though the concept itself is multifaceted (Colombetti, 2009; Shuman, Sander, & Scherer, 2013);
2. the direction of motivation, which broadly relates to the tendency of approaching or avoiding a given situation;
3. the desirability or goal-conduciveness of an emotion-eliciting situation, which focuses on the interplay between the situation and what the person is aiming to achieve.

The polarity of an affective phenomenon may therefore change greatly depending on the adopted perspective. For instance, D'Mello and colleagues (D'Mello et al., 2014) induced confusion in participants by providing, through a virtual agent in a learning environment, contradictory and incorrect information about the topic at hand. The authors found evidence suggesting that – if appropriately induced, regulated and resolved – confusion may be beneficial for learning. Is therefore *confusion* positive or negative? If we consider the first criteria proposed by Shiota and colleagues (*ibid*), confusion is usually considered unpleasant, and therefore negative on valence. The situation is nevertheless less clear-cut for the other criteria: confusion may be considered *approachable* in situations such as enigmas, and even goal-conducive in situations where the initial confusion evolves in understanding, as in the case of D'Mello and colleagues (*ibid.*).

Rowe & Fitness (2018) conducted a qualitative study on *negative* emotions in university's teachers and students, and also found mixing results. According to their analysis, negative emotions promote learning by inciting learners to seek assistance; enhancing cognition; increasing motivational drive; and enhancing productivity or performance. On the other hand, negative emotions inhibits learning by hindering communication; preventing engagement and reducing motivation; impairing cognition; and diminishing productivity or performance.

The overall picture that emerges is necessarily scattered and sometimes contradicting. Affect and learning are two complex phenomena and the combination of the two leads, exponentially, to further complexity. Considering a *simple* directional effect between intrinsic *positive* emotion and *positive* learning outcomes, and conversely between intrinsic *negative* emotion and *negative* learning outcomes, appears diminishing of both learning and affect (Hascher, 2010; Pekrun & Linnenbrink-Garcia, 2014a; Shuman & Scherer, 2014). As summarized by Pekrun and Perry (2014, p. 134) "with few exceptions, any emotion can prove to be either adaptive or maladaptive". A similar view is taken by Hascher (2010):

> It is too simple to suppose that negative emotions have negative effects on learning, and positive emotions have positive effects. What arguments can be found to clarify a statement like 'Happiness during learning is negative for the learning process'? The question of negative effects of positive emotions (and of positive effects of negative emotions) is very interesting, because it goes beyond plausibility arguments.
> — Hascher (2010), p. 15

In this regard, an EAT takes in somehow an agnostic position, since it may register and convey all sort of affective phenomena that are considered potentially relevant to the learning context in which it is deployed, regardless of whether they are supposed to hinder or foster learning. At the same time, though, the EAT should build emotional meaning making and awareness on the ground of affective phenomena that are known to play a prominent role in learning processes and outcomes.

1.1.3 Socio-Affective Competences in Learning

Finally, a third approach in research linking affect and learning focuses on empowering learners with skills, allowing them to deal with the social and emotional challenges that they face during various phases of the learning process. Oftentimes, the scope of this line of research goes beyond learning and educational settings, aiming at preparing learners – especially in early educational programs – to take advantage from their social and emotional skills in life more generally. Research of this third kind often – but not always – refers to the concept of Social and Emotional Learning (SEL), which has a long tradition in education (see Osher et al., 2016). Brackett and colleagues (2019) provide a recent definition of SEL:

> Today, SEL refers to the process of integrating cognition, emotion, and behavior into teaching and learning such that adults and children build self- and social awareness skills, learn to manage their own and others' emotions and behavior,

> make responsible decisions, and build positive relationships.
> — Brackett et al. (2019), p. 144

One extensive perspective on this form of social and emotional learning is represented by the RULER approach (Brackett et al., 2019; Brackett, Rivers, Reyes, & Salovey, 2012a; Hoffmann, Brackett, Bailey, & Willner, 2020; Nathanson, Rivers, Flynn, & Brackett, 2016), an acronym that stands for:

- Recognizing emotions in oneself and others;
- Understanding the causes and consequences of emotions;
- Labeling emotions with a nuanced vocabulary;
- Expressing emotions in accordance with cultural norms and social context;
- Regulating emotions with helpful strategies.

The RULER approach relates to various psychological theories, among which the concept of emotional intelligence as originally posited by Salovey and Mayer (1990) and more recently updated in light of further development (Mayer, Caruso, & Salovey, 2016). According to the authors, emotional intelligence is an ability comprising four branches: (a) perceive emotions accurately, (b) use emotions to accurately facilitate thought, (c) understand emotions and emotional meanings, and (d) manage emotions in themselves and others.

Other models and conceptualization of emotional intelligence have emerged in the last few decades, either under the same or a different name (Cherniss, 2010; Hoemann et al., 2021; Locke, 2005; Murphy & Hall, 2011; Scherer, 2007). In a recent scoping review conducted by Hoemann and colleagues (2021), the authors group these concepts under the banner of *emotional expertise* under the premise that they "share the observation that some people are better than others at a range of competencies related to understanding and experiencing emotions, and these competencies may help them lead healthier lives" (*ibid.*, p. 1160). From a more practical standpoint, there is nowadays a growing consensus in distinguishing between two different approaches to emotional intelligence or related formulations: one considers it a relative stable trait (*e.g.*, Petrides, Frederickson, & Furnham, 2004), the other an ability or competence that can be trained (*e.g.*, Mayer et al., 2016; Scherer, 2007) and also measured from a performance standpoint (Schlegel & Mortillaro, 2018). This second conceptualization is more keen to the learning context, assuming that it is possible to act upon emotional expertise and make it evolve over time (Brackett et al., 2019; Hoffmann et al., 2020; Scherer, 2007).

Another line of inquiry bridging learning and affect relates to various forms of regulatory processes, which are considered fundamental in learning both at the individual and collective levels. A consistent body of research, for instance, refers to three types of regulatory processes: (1) self-regulated learning, (2) co-regulated learning, and (3) socially-shared regulation of learning (Järvelä et al., 2015, 2016; Miller & Hadwin, 2015; Winne, 2015). According to Miller and Hadwin (2015, p. 574), regulation of learning "can be defined as an intentional, goal directed metacognitive activity in which learners and groups take strategic control of their actions (behaviour), thinking (cognitive), and beliefs (motivation, and emotions) in the context of dynamic social interactions". In self-regulated learning, the person intentionally acts upon her own internal states that may have effect on her learning process. In co-regulated learning, learners engage in mutual, but still individually-oriented, regulation of internal states. Finally, in socially-shared regulation, which emerges from the coordination of self- and co-regulated learning, the group as a whole becomes the target of the regulatory processes in a synchronized and productive manner (*ibid*).

This third line of inquiry, thus, consider affect in the light of integrated processes, which require both effort and *some* form of expertise for them to be instrumental to learning. This perspective resonates with the voluntary use of an EAT, which requires effort and inferential processes, but also with the possibility that the use of the EAT may contribute to better regulate socio-affective dynamics.

1.1.4 Synthesis

This first section of the chapter presented a brief overview of three complementary research areas linking affect and learning. The implementation of an EAT relates to all three areas since it presupposes (1) that learning elicits affective phenomena, so that learners have affect-related information upon which they can reflect and/or communicate; (2) that affective phenomena experienced during learning have an impact on learning processes and outcomes depending on the specifics of the situation at hand, rather then on a predefined list of discrete affective states that always foster or hinder learning ; and (3) that for taking full advantage of the affect-related information, both at the intra- and inter-personal levels, some effort and expertise in conveying meaning with, and extrapolating meaning from emotional information is necessary.

1.2 Affect-Aware Systems in Computer Mediated Learning Environments

The pivotal role attributed to affective phenomena in learning and education at large has recently pushed several scholars to investigate how to endow computer-mediated learning environments with systems that integrate affective phenomena (Arguedas et al., 2015; Calvo & D'Mello, 2010, 2012; Calvo et al., 2015; Cernea & Kerren, 2015; Feidakis, 2016; Grawemeyer et al., 2017; Harley et al., 2017). These systems, which are often referred to as affect-aware or emotion-aware systems (*ibid.*), may vary greatly in their complexity, scope, ways to measure affective phenomena, and use of the affective information. For instance, affective systems can be oriented – exclusively, primarily or equally – towards the learner's individual affective states, the collective affective states of a group sharing common learning processes and outcomes, the affective states of teachers, or even the affective states of computerized agents implemented into the system, such as embodied tutors (Cernea et al., 2014; Craig, D'Mello, Witherspoon, & Graesser, 2008; Lavoué et al., 2020; M. Lee et al., 2016; Mänty, Järvenoja, & Törmänen, 2020; Näykki, Järvelä, Kirschner, & Järvenoja, 2014).

Even though affective systems are mainly driven by autonomic affect detection (Calvo & D'Mello, 2010) and often aim at providing the learning environment with some form of adaptive intelligence, which reacts to or acts upon learners affective states (D'Mello &

Graesser, 2015; Harley et al., 2017), recent and overarching contributions have attempted to put affective systems in computer-mediated learning environments into broader perspectives (Calvo & D'Mello, 2012; Cernea & Kerren, 2015; D'Mello & Graesser, 2015; Feidakis, 2016; Harley et al., 2017). It is therefore worth to take advantage of the extant literature to seek for commonalities and differences with respect to an EAT with the characteristics stated in the introduction.

1.2.1 Affective Systems and Affective Computing

Affective systems can be broadly categorized according to the objectives of affective computing, an interdisciplinary field investigating "computing that relates to, arises from, or deliberately influences emotions" (Picard, 2000, p. 3). Picard (2009) identifies four non mutually exclusive research areas in affective computing:

1. technologies for sensing, recognizing, modeling, and predicting emotional and affective states;
2. methods for computers to respond intelligently and respectfully to handle perceived affective information;
3. technology for displaying emotional information or mediating the expression or communication of emotion;
4. computational mechanisms that stimulate internal emotions or implement their regulatory and biasing functions.

Affective systems in computer-mediated learning environments more often relate to the first two categories identified by Picard (Calvo & D'Mello, 2012; D'Mello & Graesser, 2015; Feidakis, 2016; Harley et al., 2017). For instance, in their glossary to an overarching chapter on the integration of affect in advanced learning technologies, D'Mello and Graesser (D'Mello & Graesser, 2015, p. 2) define an affect-aware learning technology as "an intelligent technology that considers a learner's affective and cognitive states in its pedagogical decision making". The authors (*ibid.*) identify two ways in which affect-aware learning technologies manifest their *intelligence*: (1) by proactively inducing or impeding particular affective states, which are considered as either fostering or hindering learning; and (2) by reacting to specific affective states as they arise, for instance in an attempt to remedy to affective states that may undermine the overall learning experience. More recently, Harley and colleagues (Harley et al., 2017) have extended the dichotomy proposed by D'Mello and Graesser (*ibid.*) and adopted it in an effort to provide a theoretically-guided taxonomy for the development of emotion-aware learning technologies, which foster *positive* emotions. The taxonomy itself is not relevant for the present contribution, since it presupposes some form of *intelligence* in the system. It is nevertheless worth mentioning Harley and colleagues (*ibid.*) contribution, for it adopts Pekrun's Control-Value theory of achievement emotions (Pekrun, 2006), already introduced above, as the guiding framework for the taxonomy, corroborating the importance to bridge techno-centric and theoretical approaches.

1.2.2 An Interactional Approach to Affect[1]

At first blush, the third research area of affective computing identified by Picard (2009) stands out as the best match for an EAT: technology for displaying emotional information or mediating the expression or communication of emotion. According to Boehner and colleagues (2007), though, this particular line of research has been neglected in affective computing, whose mainstream approach limits emotion to an internal and individual phenomenon. They identify this mainstream approach to emotion as *informational*, which they qualify as "rooted in a longer laboratory-science tradition of studying emotion in which subjective experience is considered suspect, to be replaced by objective measures" (Boehner et al., 2007, p. 276). The authors advocate instead an *interactional* approach, which (1) focus on how computers can help users understand their emotions, rather than how computers can understand users' emotions; and (2) sees emotions as a cultural, dynamic and social phenomena, constructed in action and interaction. The position taken by Boehener and colleagues (*ibid.*) is stated very clearly in this passage:

> The role of affective systems is not to transmit pre-existing emotional units, but to provide a resource for emotional meaning-making. Success of such a system is measured by whether users find the system's responses useful for interpreting, reflecting on, and experiencing their emotions. Evaluation does not aim at finding a user's original, "true" emotions, but in tracking how emotions are constructed and interpreted over time, and correlating these dynamics with aspects of system design.
> — Boehner et al. (2007), p. 287

More recently, Cernea and Kerren (2015), inspired by the work of Boehener and colleagues (*ibid.*), proposed a categorization of technologies *on the rise* for emotion-enhanced interaction. The authors focused "on technologies that satisfy two main requirements: having been used increasingly often in affect detection for interaction purposes, and having a proven track of *portability*, *non-intrusiveness* and *low cost*" (Cernea & Kerren, 2015, p. 72, italics in the text). Even though the authors target more complex affective system – often using multi-modal sensory information, such as tracking and recognizing facial expressions, eye-movements or pupil dilation, or brain-computer-interfaces – the categorization they propose is relevant also to a *simpler* system. The authors identify three types of affective systems from an interactional stand-point, proposed here in an condensed form (*ibid.*, p. 78):

1. *Emotion-adaptive systems*, which adapt interaction based on the user's affective state *or* mimic affective states with the purpose of enhancing the user experience and efficiency;
2. *Emotion modeling systems*, which do not simply adapt to user affective states, but try to influence these states as well;

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

3. *Affective-awareness systems*, designed to improve affective self-awareness of a user (i.e., internal affective awareness) *or* awareness of the emotions of other users (i.e., external affective awareness).

An EAT belongs clearly to the last category of *affective-awareness systems.* The presence of this category corroborates the fact that emotional awareness is considered crucial also from an affective computing perspective.

1.2.3 Emotion-Aware Systems in E-Learning

Examples of contribution from an *affective-awareness systems* perspective mainly driven by affective computing is provided by Feidakis and colleagues (Feidakis, 2016; Feidakis et al., 2014; Feidakis et al., 2011). The authors investigated emotion-aware systems in computer-mediated environments, both from theoretical and empirical perspectives, highlighting though that the topic has received so far limited attention. Their work is also primarily concerned with intelligent learning environments, but some of their contribution have more similarities with the use of an EAT as posited in the present thesis.

For instance, the authors draw extensively from emotion theories (Feidakis, 2016; Feidakis et al., 2014). Furthermore, in an overarching review of emotion-aware systems for e-learning, Feidakis and colleagues (2016, p. 217) define emotion awareness as "the *implicit* or *explicit collection of emotion data* and the *recognition* of *emotion patterns*" (italics in the text). They therefore draw a distinction between data derived from autonomic affect detection (implicit), and data provided through first-person subjective report of feelings (explicit). Feidakis and colleagues' work is also relevant to the present contribution because they advocate the adoption of emotion-aware systems into *simpler* learning management systems, whereas affective systems are often deployed in more technology-intensive settings, such as virtual environments or educational games.

Furthermore, affect- or emotion-aware systems often – but not always – complement affective detection with what are commonly referred to as *affective feedbacks* (Arguedas et al., 2015; Feidakis et al., 2014; Grawemeyer et al., 2017; Robison, McQuiggan, & Lester, 2009), which are distinguished from more traditional *cognitive feedbacks* (M. M. Nelson & Schunn, 2009). As the field more generally, affective feedbacks in affective systems are a recent line of inquire and the extant literature is therefore very cautious about their efficacy, as well as the very nature of what represents an affective feedback (*ibid.*). Broadly speaking, an affective feedback is a reaction of the system that explicitly targets the affective phenomena detected and attempts to instrumentally act upon it to improve learners' situation. The reaction may be driven either from an affective or task-oriented point of view (Robison et al., 2009). In the first case, the system shows directly affective reactions itself, for instance by showing *empathetic* reactions, or try to regulate specifically learners' affective states. In the second case, the affective information is integrated directly in to the instructional design, for instance by adapting the difficulty of the task.

Affective feedbacks are tightly related to a form of *intelligence* in affective systems, which computes appropriate treatment and useful output, and is therefore outside the scope of a simpler system. They nevertheless call for attention in evaluating whether emotional expression can be effective on itself, or it should be coupled with some form of feedback from the system.

1.2.4 Self-Report From An Affective Computing Perspective

A last element of interest that can be derived from the literature driven mainly by an affective computing perspective concerns how self-report is assessed. Given that this field makes extensive use of autonomic identification of affective phenomena, it is of interest to seek for drawbacks in the use of a voluntary, self-report tool. Views from an affective computing perspective can therefore complement, with information derived from computer-mediated environment, more *traditional* assessment about validity and reliability of self-report in measuring affective phenomena (*e.g.*, Mauss & Robinson, 2009; Mortillaro & Mehu, 2015; Scherer, 2005).

In an overarching contribution about measuring emotions in computer-mediated learning environments – referred to as Computer-Based Learning Environments (CBLE) – Harley (Harley, 2015) reviews four main methodologies: facial expression coding, body posture and physiological measurement devices, log file data, and self-report measures. The author confirms that self-report is a widely adopted method of measure even in CBLE, which can be administered before, during or after the learning activities. It has the advantage of being easy to administer, and collected data are generally also easy to analyze and compare, especially when the self-report consists of a set of discrete affective states. On the other hand, though, Harley (*ibid.*) points out that self-report is an *off-line* measure, since learners must divert their attention from the learning activity to the self-report measure. This diversion is particularly disrupting when the self-report is asked at frequent intervals, or is repetitive in its administration. These seem the main drawbacks that are specific, or at least of particular relevance, to CBLE. Other drawbacks identified in the contribution are in fact shared with general and known shortcomings in self-report measures more generally, and can be roughly divided in: (a) authenticity of reported experience; (b) unwillingness to disclose one's own affective experience for normative pressure or social desirability; (c) latency time between the felt affective phenomena and the report, which may introduce bias in recollection; and (d) the action of self-reporting may elicit corollary phenomenon that modifies the original affective state (Harley, 2015; Mauss & Robinson, 2009; Mortillaro & Mehu, 2015; Scherer, 2005).

1.2.5 Synthesis

This section provided an overview of affective systems implemented in computer-mediated learning environments from perspectives mainly driven by affective computing. The literature provides similarities between affective systems and an EAT as the one prospected in the thesis. First, there is a growing interest in affective computing for *interactional* affective technologies that provide users with the possibility to impinge on emotion meaning-making. Second, emotional awareness is an active research topic in affective computing, and in computer-mediated learning environments more specifically. Third, even though the field of affective computing was originally theory-agnostic and more focused on technical features in affect detection and processing (Calvo & D'Mello, 2010), recent developments in the field corroborate a more tight relationship with theories, which can be beneficial in both direction: theory can inform affective systems, and affective systems can inform theory (Calvo & D'Mello, 2010; Marsella et al., 2010a; Pekrun & Linnenbrink-Garcia, 2014a; Scherer et al., 2010).

On the other hand, most affective systems make the assumption that specific affective phenomena can hinder or foster learning and, congruently, deploy proactive or reactive strategies to intervene in the process. As stated in the introduction, the kind of EAT targeted by the thesis attempts to maintain a certain degree of agnosticism towards the instrumentality of specific affective phenomena and bestow learners with the responsibility to integrate the emotional information at hand. In this sense, an EAT is therefore closer to an awareness tool than it is to an affective system. The next section thus proposes an overview of the literature on awareness tools.

1.3 Awareness Tools in Computer-Mediated Learning Environments

Whereas affective systems in computer-mediated learning environments stem directly from affective computing, awareness tools adopted in learning settings mainly originated from research in Computer-Mediated Communication (Baltes et al., 2002; Fjermestad, 2004; Licklider & Taylor, 1968) and Computer-Supported Cooperative (or Collaborative) Work (Dourish & Bellotti, 1992; T. Gross, Stary, & Totter, 2005; Grudin, 1994). Consequently, they are essentially driven by an inter-personal perspective, complying with the need to provide users sharing a computer-mediated environment with instrumental information about others to perform the task at hand (Dourish & Bellotti, 1992; Gutwin & Greenberg, 2002; Gutwin, Greenberg, & Roseman, 1996; Schmidt, 2002). The concept has later been integrated in computer-mediated learning environments from different perspectives and with different approaches (Buder, 2011; Engelmann et al., 2009; Janssen et al., 2011; Jézégou, 2010; Kirschner et al., 2015; Lowenthal & Snelson, 2017; Tu & McIsaac, 2002). As a result, the concept of *awareness* covers a wide range of intra- and inter-subjective processes that may sustain learning at the cognitive, social and affective levels, and also target one or more form or regulation: self-regulation, co-regulation, and socially shared regulation (Järvelä et al., 2015, 2016; Miller & Hadwin, 2015; Winne, 2015).

This section starts with an overview of the flexible notion of *awareness* in computer-mediated learning environments, starting by a functional definition of the concept. Next, it focuses on awareness tools more specifically, by highlighting how current trends in the interpretation of the role of these tools are compatible with emotional awareness. The section concludes with two concepts widely adopted in computer-mediated learning environments, which may benefit from emotional awareness: social presence in distance learning, and mutual-modeling in Computer-Supported Collaborative Learning.

1.3.1 Awareness Results from Displaying and Monitoring Functions

As stated in the preamble of this section, the term awareness in the context of awareness tools emerged primarily from the fields of Computer-Mediated Collaboration (CMC) and Computer-Supported Cooperative Work (CSCW) as a means to overcome reduced contextual information in computer-mediated environment with respect to face-to-face interaction (Bodemer & Dehler, 2011; Buder, 2011; T. Gross et al., 2005; Schmidt, 2002). Even within CMC and CSCW, though, the concept has been used with different meanings, with often various preceding adjectives that attempted to better qualify the phenomenon such as *general awareness*, *collaboration awareness*, *peripheral awareness*, *background awareness*, *passive awareness*, *reciprocal awareness*, *mutual awareness*, or *workspace awareness* (Schmidt, 2002, pp. 286–287 citing use in other articles).

Schmidt (2002) argues that such attempts are doomed to fail, because awareness is not a distinct mental state that exists *per se*. It rather refers to being aware about something, and it is *that* something that determines what awareness is. In this regard, there are two intertwined functions that allow awareness to emerge. First, information must be made available, or produced, for awareness to be possible. Schmidt (*ibid.*) define this function as *displaying*. The emergence is nevertheless not enough, information must also be perceived, or consumed, so Schmidt define a second necessary function, which consists in *monitoring* the computer-mediated environment in search of displayed information. As a consequence, awareness "is not the product of passively acquired 'information' but is a characterization of some highly active and highly skilled practices" (Schmidt, 2002, p. 293). In the reminder of the section and the manuscript, awareness is therefore considered in functional terms as the result from the interaction of displaying and monitoring instrumental information.

1.3.2 Instrumental Information in Computer-Mediated Learning Environments

Defining awareness as the function of displaying and monitoring instrumental information determines the need to frame what is considered as instrumental information. In this regard, when the concept of awareness was transferred into the context of learning, a change of perspective materialized on some important aspects, which are here resumed in three separated points, even though there is ample overlapping and interaction among them.

First, there is a growing consensus in considering that awareness in computer-mediated learning environment should not attempt to reproduce the face-to-face golden standard (Bodemer & Dehler, 2011; Buder, 2011; Janssen et al., 2011; Kirschner et al., 2015). This perspective is warranted by two intertwined assumptions.

On the one hand, the *universal* primacy of face-to-face over computer-mediated interactions has been challenged by several scholars (Baltes et al., 2002; Buder, 2011; Buder, Bodemer, & Ogata, 2021; Derks et al., 2008; Fjermestad, 2004; Stone & Posey, 2008). Even though the absence of para-verbal cues that are usually available in face-to-face interactions may impoverish inter-personal dynamics, endowing computer-mediated learning environment with seamless audio/video connection is not the solution to all problems (*ibid.*), especially in asynchronous situations. One integral element of face-to-face interactions consists in the social exchanges that can sustain learning processes, even in informal or auxiliary forms (Kreijns et al., 2003, 2013). Kirschner and colleagues (2013), for instance, provide the example of the water cooler as a *social affordance* (*ibid.*), which prompts learners to gather and talk about various aspects of their learning experience. According to the authors, thus, computer-mediated learning environments should not only provide the equivalent of the water cooler, but also ensure that learners use it, since the mere availability of social affordances does not assure neither their use, nor their effectiveness (Kreijns et al., 2003).

On the other hand, information and communication technologies available in computer-mediated environments allow learners and instructional designers to provide information that is not necessarily available – or at least hard to notice, process or remember – in face-to-face interaction (Buder, 2011; Buder et al., 2021). In other words, rather then reproduce the contextual information available in face-to-face interaction, instructional designer should endow computer-mediated learning environments with content, activities, or technological artifacts that leverage on computational and human-computer interaction principles to improve learning processes and outcomes. For instance, learners could display their background knowledge about the topic at hand and monitor previous knowledge of their colleagues (Dehler, Bodemer, Buder, & Hesse, 2011; Mirweis Sangin et al., 2011), reducing the difficulty in judging what someone knows about a subject (Nickerson, 1999, 2001). Another possibility provided by computer-mediated environment consists in sharing information about learners' levels of participation in collaborative or cooperative settings (Janssen et al., 2011), which may contribute to reduce free-riding and social loafing behaviors (Karau & Williams, 1993; Salas, Grossman, Hughes, & Coultas, 2015).

Second, even though awareness is considered a secondary – or corollary – source of information with respect to the *content* of the learning activity (Janssen & Bodemer, 2013), scholars are considering a more flexible approach to the amount of resources that can or should be oriented to displaying and monitoring instrumental information that is not directly related to the learning outcome (Bodemer & Dehler, 2011; Dillenbourg et al., 2016; Kreijns et al., 2003). Rather than minimizing at all costs exogenous activities in an attempt to limit distraction and interference caused by a dual-task (Pashler, 1994), researchers are contemplating trade-offs based on the ratio between the effort of displaying and monitoring information, and the benefit resulting from disposing of that information (Dillenbourg et al., 2016). This trade-off is particularly important since researchers agree that awareness seldom emerges as a natural by-product of the learning task *per se*, but rather needs to be appropriately enhanced (Bodemer & Dehler, 2011; Kreijns et al., 2003). In this regard, two non-mutually exclusive pedagogical strategies are often identified in the literature as a means to enact awareness: *scripting tools* and *group awareness tools* (*e.g.*, Miller & Hadwin, 2015).

Scripting tools (*e.g.*, Dillenbourg, 2002; F. Fischer et al., 2013) are mainly adopted to scaffold the learning process, so that awareness is created as the results of planned interactions. In other words, learning is organized as a sequence of steps, where some or all of these steps foster or require the emergence of awareness, by explicitly asking learners to display and monitor instrumental information. A simple – and questionable – example of this process are the many cases of Massive Open Online Courses (MOOCs), which prompt thousands of users, located all over the world, to present themselves to others, often by posting a message in a forum with information about themselves and their learning goals. More elaborate scripts may require learners to engage in collaborative or argumentative activities, negotiation of goals and means to attain them, or discussion between peers about the quality of productions or interactions (*ibid*). The common factor shared by all the possible activities is the fact that awareness is mainly produced by the *script*, that is, by the learning design itself.

Group awareness tools (*e.g.*, Buder, 2011; Buder et al., 2021; Engelmann et al., 2009; Janssen & Bodemer, 2013), often shortened in awareness tools, on the other hand, take a different and mainly unstructured perspective. Rather than sequencing specific moments or activities during which awareness is enacted, group awareness tools bestow learners with the responsibility of displaying and monitoring information whenever it is considered appropriate and instrumental to the learning activity (*ibid.*). In other words, group awareness tools are usually integrated alongside the learning activity – persistently or *on demand* – and support learners in making others aware of one's instrumental information, as well as becoming aware of others' instrumental information. As stated by Miller and Hadwin (2015. p. 582), compared to scripting tools, "group awareness tools take a more non-directive or reactive approach placing the locus of control in the hands of the learners". In this regard, Buder (2011) – adopting Schmidt (2002) functional definition of awareness in learning contexts – identifies eight features of awareness tools about which the trade-off between effort and benefit of bestowing learners with the locus of control should be assessed. For the displaying function, Buder (*ibid*) identifies the following four features:

- To what extent information is displayed explicitly or implicitly, with explicit information demanding more effort and therefore potentially interfering more with ongoing learning activity, but also providing more meaningful and reflected-upon information;
- The frequency by which new information is provided, which is inherently linked with the previous feature insofar as the information is displayed explicitly;
- To what extent information is displayed voluntarily or is enforced by the system, for instance by blocking other interactions with the computer-mediated environment until the information is not provided;
- To what extent information is displayed in a closed (*i.e.*, structured) or open (*i.e.*, unstructured) format, where closed formats usually foster immediacy but limit freedom, for instance by clicking a pre-determined choice, whereas open formats allow more flexibility.

With respect to the monitoring function, Buder (*ibid.*) identifies four other features on which trade-offs in noticing and processing the available information should be assessed. These features are the following:

- To what extent available information is integrated with the learning content or activity, assuming that, in general, the more distant the information, the greater the effort necessary to switch between the primary learning task and auxiliary awareness;
- To what extent available *chunks* of information are easy or hard to process. For instance, information may be aggregated or graphically represented to foster compar-

ison between information belonging to the learner herself and information provided by other participants to the computer-mediated learning environment;
- To what extent available information foster or pressure the display of further information. Learners may benefit from the social affordance provided by the tool, which reminds them of the interest to share instrumental information; but may also feel coerced to share information by normative pressure, for instance in an attempt to match the amount of information provided by others;
- To what extent available information prompts immediate or diffuse consequences on learners behavior. Immediate consequences may be of proximal instrumentality, but also disrupt the course of action; whereas diffuse consequences are more flexible, but may require more meta-cognitive and reflective efforts.

Third, even though there is a tendency to divide group awareness tools in three categories – *behavioral awareness*, *cognitive awareness*, or *social awareness* (Bodemer & Dehler, 2011) – depending on the information they provide, there is also an increasing consensus in considering these categories as non-mutually exclusive, allowing interactions or spill-over effects between *types* of awareness (Buder, 2011; Janssen & Bodemer, 2013; Kirschner et al., 2015). Furthermore, social and affective dimensions are receiving increasing attention in an attempt to compensate "the tendency to restrict social interaction to educational interventions aimed at cognitive processes while social (psychological) interventions aimed at socio-emotional processes are ignored, neglected or forgotten" (Kreijns et al., 2003, p. 336). In this regard, the next two subsections provide examples of theories that integrate the concept of awareness from cognitive, social, and affective perspectives in computer-mediated learning contexts: social presence in distance learning, and mutual-modeling in Computer-Supported Collaborative Learning.

1.3.3 Social Presence in Distance Learning

The broad concept of awareness can be related with issues that have been identified in the last few decades about the articulation between distance learning and the need for *social presence*, especially in higher-education (A. W. Bates & Bates, 2005; Garrison, Anderson, & Archer, 2010; Jacquinot, 1993; Jézégou, 2010; Kreijns et al., 2003; Rourke, Anderson, Garrison, & Archer, 2001; Tu & McIsaac, 2002). The definition of social presence and whether this presence is a necessary element for learning to happen is still source of lively debate (Jézégou, 2010; Kirschner et al., 2015; Lowenthal & Snelson, 2017; Rourke & Kanuka, 2009). For instance, Lowenthal and Snelson (2017) recently collected definitions from the most cited articles on social presence and identified five categories in which different definitions could be divided into (*ibid.*, see Table 1 in the original article, p. 144):

- Social presence as *being there*, as in the definition by Dunlap and Lowenthal (2009): "the degree of salience (i.e., quality or state of 'being there') between two communicators";
- Social presence as *being real*, as in the definition by Gunawardena and Zittle (1997): "the degree to which a person is perceived as 'real' in mediated communication";
- Social presence as *projecting*, as in the definition by Rourke and colleagues (1999):

"the ability of learners to project themselves socially and affectively into a community of inquiry";

- Social presence as *connecting*, as in the definition by Tu (2002): "the degree of feeling, perception, and reaction of being connected on CMC [computer-mediated communication] to another intellectual entity";
- Social presence as *belonging*, as in the definition by Picciano (2002): "a student's sense of being in and belonging in a course and the ability to interact with other students and an instructor".

Another definition of particular interest to the present contribution is provided by Kreijins and colleagues (2011, p. 366), who define social presence "as the degree of illusion that others appear to be a 'real' physical persons in either an immediate (i.e., real time/synchronous) or a delayed (i.e., time-deferred/asynchronous) communication episode". This definition explicitly refers to the synchronous and asynchronous temporal dimensions, which are of interest in the empirical contributions in Part III.

The debate about which category or specific definition better defines social presence, though, is beyond the scope of the present contribution, for which rather than the differences, it is worth pointing out commonalities, especially in considering socio-affective phenomena. In this regard, the definition of Rourke and colleagues (1999) clearly cites affect as a necessary dimension for learners' to project themselves in the social *milieu*, in which learning happens. Vaughan and Garrison (2005), cited in Lowenthal & Snelson (2017), states it very similarly: "social presence refers to the potential of participants to project themselves socially and emotionally". A view shared by Swan and colleagues (2008), also cited in Lowenthal & Snelson (2017), for whom "social presence refers to the degree to which learners feel socially and emotionally connected with others in an online environment". In the work of Jézégou (2010), *socio-affective presence* is one of the three fundamental dimensions of social presence alongside *cognitive presence* and *pedagogical presence*. According to Jézégou (*ibid*):

> the socio-affective presence in e-learning allows to support the cognitive presence resulting from learners' transactions to solve a problematic situation. It is generated by social interactions that contribute to establishing the symmetry of the relationship and the friendliness between learners, thus creating a socio-affective climate that is favorable to transactions within a digital communication space.
> — Jézégou (2010), p. 267, our translation from French.

There is therefore a growing consensus in considering affective phenomena as an integral part of social presence in distance learning. Displaying and monitoring information about learners' affective states can therefore contribute to enhance the social presence , by providing learners with the possibility to project themselves socially and affectively in computer-mediated learning environments.

1.3.4 Mutual Modeling in Computer-Supported Collaborative Learning

Computer-Supported Collaborative Learning (CSCL) is another inter-disciplinary domain of research adopting awareness as a focal concept is(Cress, Rosé, Wise, & Oshima, 2021;

Dillenbourg, 1999; Dillenbourg et al., 2009; Stahl et al., 2006; Suthers, 2006). The field of CSCL – which is not limited to remote learning but encompasses also co-located interaction – has, from the very beginning in the 1990's, reckoned the importance for learners to build a *shared-understanding* of the learning activity (Cress et al., 2021; Dillenbourg et al., 2009, 2016; Roschelle & Teasley, 1995). For instance, Suthers (2006, p. 332) posits that CSCL is characterized by "processes of intersubjective meaning making and how technological affordances mediate or support such processes". Technology affordances are defined by Norman (2013, p. 11) as the "relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used" (see also Pucillo & Cascini, 2014; Rizzo, 2006; Turner, 2005). In other words, the field stresses the importance of technology as a support for learners to build their learning processes from the *bottom-up*, rather than having knowledge passively delivered to them. In this sense, the shared understanding and the inter-subjective meaning-making are necessary, even though not sufficient, conditions for planning, initiating, maintaining, negotiating and regulating those processes (Janssen & Bodemer, 2013; Järvelä et al., 2015, 2016; Kirschner, Strijbos, Kreijns, & Beers, 2004; Kreijns et al., 2003, 2013).

One way to build and maintain shared understanding and inter-subjective meaning making has been identified as partner- and mutual-modeling (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; Mirweis Sangin et al., 2011). In fact, according to Dillenbourg and colleagues (2016), learners must engage in two intertwined processes: (1) a learner has to build and maintain a holistic representation of the colleagues with whom she shares the learning activity; and (2) the very same colleagues must possess sufficient information to build and update a holistic representation of the learner herself. Dillenbourg and colleagues define the former process as *partner modeling* – that is "the process of inferring one's partner's mental states" (*ibid*, p. 230) – and the latter as *mutual modeling*, that is, *bi-directional* partner-modeling, which may happen at various nested levels (*e.g.* Jane speculates that Paul is aware of the fact that she expects him to be more involved in the argumentation).

It is posited, in fact, that the effort learners put into building a symmetrical representation of the partners in a collaborative task is a pivotal determinant of learning processes and, by extension, outcomes (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; Mirweis Sangin et al., 2011). In this regard, though, it is useful to point out some important elements.

First, symmetry does not mean that all implicated learners must share exactly the same mental states. On the contrary, collaborative efforts are inherently characterized by fluctuations between socio-cognitive tensions and relaxations, which are instrumental to the learning process (Andriessen et al., 2011; Järvelä et al., 2016; Winne, 2015).

Second, unlike other similar concepts implicating interpersonal accuracy (see Hall et al., 2018 for an integrative overview), mutual-modeling does not have to be persistently and precisely accurate: that would require an effort that may divert resources from the learning activity (Dillenbourg et al., 2016). Mutual-modeling is therefore particularly important at times when there are events of major concern for learners, who may benefit from cues about the suitability of persisting with or changing their behavior. As we will see later, this mechanism incidentally shares many commonalities with eliciting events in emotions (Scherer, 2005). Finally, even though mutual-modeling is mainly driven from a communicative point of view (Clark & Brennan, 1991), it also encompasses a broader variety of sources of information, such as what a person does, thinks, wants, or – precisely – feels about herself or about their colleagues (Bodemer & Dehler, 2011; Dillenbourg et al., 2016; Eligio et al., 2012; Engelmann et al., 2009; Kirschner et al., 2015; Mirweis Sangin et al., 2011).

An example of this integrated perspective is provided by Janssen and Bodemer (Janssen & Bodemer, 2013), which argue against a clear-cut separation between cognitive and social awareness. In their framework, illustrated in Figure 1.1, the authors suggest that there is inter-dependency between cognitive and social information on the one side, and the content and relational space in which the collaboration takes place on the other side. As it will more thoroughly illustrated in Chapter 2, emotions are known to play a prominent role both cognitively and socially, and are therefore particularly suited to integrate information at different levels.

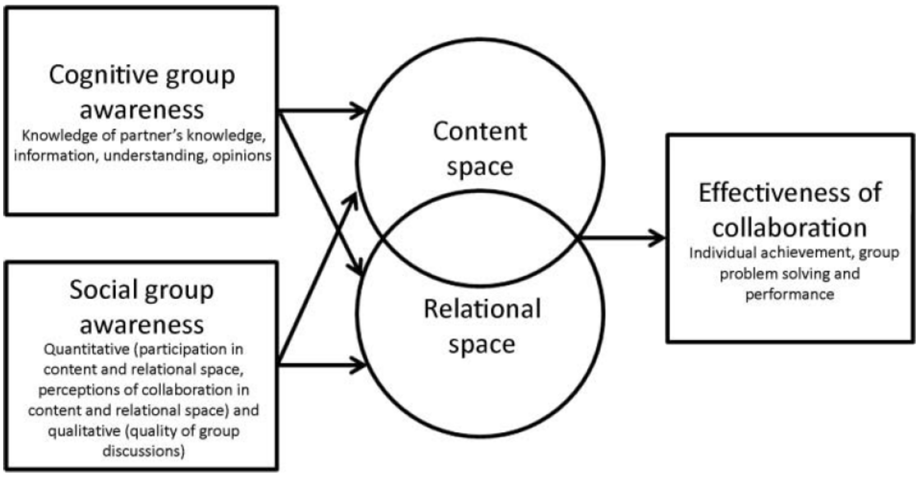


Figure 1.1: Framework of relationships between cognitive and social group awareness, coordination of the content and relational space, and effectiveness of collaboration. From the original Figure 1 in Janssen & Bodemer (2013), p. 52.

There is therefore a growing interest in exploring awareness tools that integrate into the CSCL environment information upon which learners can form a better representation of each others, in an attempt to build and maintain shared understanding and inter-subjective meaning-making. More traditional information, such as what a learner is doing in the collaborative space, are progressively integrated with information stimulating more meta-cognitive and socially oriented processes.

1.3.5 Synthesis

This section provided an overview of the literature about awareness tools in computer-mediated learning environment as a means to display and monitor instrumental information for learning processes and outcomes. The section highlighted how the flexible concept of awareness encompasses a wide variety of potential sources of instrumental information.

In this regard, there is an increasing convergence in considering affective phenomena as viable information, for instance, to build and update a social presence in remote learning environment or a holistic representation of the partners in a Computer-Supported Collaborative Learning task. The literature on awareness tools also points out that these tools bestow learners with the locus of control in making the best out of the information provided through the awareness tool. That is, it is mostly up to learners (1) to decide when to display or monitor the information; and (2) how to extrapolate meaning from the available information and integrate it in their course of action during the learning activity.

1.4 Summary

This chapter provided an overview of the interplay between affect and learning from different perspectives. In computer-mediated learning environments in particular, affective information is integrated through the use of technological artifacts, which provide learners with the possibility to express and perceive affect-related information considered to be instrumental for learning processes and outcomes. This theoretical convergence about the instrumentality of affect-related information needs a concrete, technical way to be implemented. In this regard, this chapter highlighted some features of affect-aware systems on the one hand, and of awareness tools on the other, which have contributed to define a broad sketch of some of the features an EAT should provide, as well as some of the conditions or intervening factors that may influence its use and usefulness. Compared to affective systems, awareness tools are more aligned with the type of EAT contemplated by the present contribution. In the meantime, awareness tools are mostly oriented by the idea that *others* could benefit from information provided by oneself, whereas the individual benefits of displaying information are considered to a lesser extent. As previously mentioned, affective phenomena are not often central in learning environments, and therefore emotional *self-awareness* may also play a prominent role in an EAT. These aspects are further defined in the following chapter.

Chapter 2

Emotional Awarenss and Emotional Awareness Tools in Computer-Mediated Learning Environments

After a general overview about the relationship between affect and learning, this chapter focuses on emotional awareness and on Emotion Awareness Tools (EAT) more specifically. First, the concept of emotional awareness is defined within the context of computer-mediated learning environments. Second, three main assumptions are derived by this definition. The theoretical underpinnings of each assumption are set forth, complemented by the overview of related empirical works. In the second part of the chapter, an abstract model of an EAT is laid out in an attempt to frame the different interactions between the learners and an EAT, for emotional awareness to be instrumental. The model also allows to highlight how the instrumentality of an EAT is far from being trivial and rather requires a complex interplay between learners expertise and willingness in taking emotional awareness into account on the one hand, and features of the EAT that can sustain or facilitate the process. Finally, it is suggested that the abstract model can be used as a way to direct research in emotional awareness, by pointing out which part of the model the research aims at investigating (*e.g.*, the whole or some specific parts).

2.1 Definition of Emotional Awareness in Computer-Mediated Learning Environments

Emotional awareness in computer-mediated learning environments is often proposed in the literature as self-explaining or defined somehow implicitly. For instance, as stated in the previous chapter, Feidakis and colleagues (2016, p. 217) define emotion awareness as "the *implicit* or *explicit collection of emotion data* and the *recognition* of *emotion patterns*" (italics in the text). Cernea and colleagues (2015), in their overview of rising technologies in emotion-enhanced interaction, identify the category of affective-aware systems as "designed to improve affective self-awareness of a user (i.e.,internal affective awareness) or awareness of the emotions of other users (i.e.,external affective awareness)" (*ibid.*, p.78). More precise definitions are provided by Lavoué and colleagues (2020), who partially derive them from clinical psychology. The first definition provided by Lavoué and colleagues is based upon Boden & Thompson (2015) and posits that emotional awarenss is "the ability to perceive, identify, and understand emotions" (Lavoué et al., 2020, p. 270). Later in their article, Lavoué and colleagues (*ibid.*) propose another definition, based upon Rieffe and colleagues (2008), referring to emotional awareness as "the attentional process by which individuals identify, explain and differentiate between their own emotions as well as the others' emotions" (Lavoué et al., 2020, p. 270). These two definitions resonate with socio-emotional competences briefly illustrated in the previous chapter. Finally, later in the same article, Lavoué and colleagues (*ibid.*) provide a more technical definition of emotion awareness tools as "tools that display information on own' own [sic] or partners' emotions, circumstances and antecedents" (Lavoué et al., 2020, p. 284). Compared to the more abstract definition in Feidakis and colleagues or in Cernea and colleagues (*ibid.*), the definition in Lavoué and colleagues (*ibid.*) implies that the tool shall also provide information about circumstances and antecedents that concurred in eliciting the emotional episode.

From these definitions it is possible to extrapolate three fundamental assumptions about emotional awareness and, by extension, the instrumentality of an EAT in a computer-mediated learning environment:

1. Learners may benefit from intra-personal emotional awareness by using information about their own emotions as valuable information for their own learning processes (self-regulation).
2. Learners may benefit from inter-personal emotional awareness by using information about their partners' emotions and/or by communicating their own emotions to their partners and/or by a combination of the two. One or all of these circumstances may contribute to learners' own learning processes (self-regulation), the learning processes of their partners individually (co-regulation), or the learning processes of the group as a whole (socially shared regulation).
3. Learners may benefit from emotional awareness conveyed by a dedicated tool, which *encodes* emotional information into the computer-mediated learning environment, and *decodes* that information, for learners to extrapolate emotional meaning-making instrumental to learning.

Each assumption is discussed in the following three sections. Every section proposes an overview of related theories, as well a selection of empirical works, which are of particular interest for the present contribution in terms of objectives, methodologies, or findings.

2.2 Emotional Awareness at the Intra-Personal Level

The first assumption about the instrumentality of emotional awareness implies that learners can benefit from it at the individual, intra-personal level.

2.2.1 Theoretical Underpinnings

The role of emotion at the intra-personal level has received over the last few decades extensive consideration from different perspectives and using different methods of investigation. A widespread consensus has emerged over the years about the fact that emotion, rather then disruptive of behavior and in antithesis with cognition, plays a prominent role in helping the organism to cope with a complex environment (Adolphs & Anderson, 2018; Armony & Vuilleumier, 2013; Damasio, 2006, 2018; Immordino-Yang, 2016; Levenson, 1999, 2014; Leventhal & Scherer, 1987; Pessoa, 2013; Sander, 2013; Scherer, 2009a).

For instance, emotion is known to influence high-level cognitive functions such as attention, perception, memory, and decision-making (Brosch et al., 2013), which are all implicated in learning processes. For instance, it is posited that an emotional attention system may complement and interact with the exogenous and endogenous systems (Brosch, Pourtois, Sander, & Vuilleumier, 2011). Emotionally charged stimuli are also known to often bestow precedence in perception over non emotionally-charged stimuli (Brosch, Pourtois, & Sander, 2010). Many links between emotion and memory have been established both in the encoding and retrieving of information (Kensinger & Ford, 2020; Sharot & Phelps, 2004), which may have important consequences for learning (Beege, Schneider, Nebel, Hässler, & Günter, 2018; Tyng, Amin, Saad, & Malik, 2017). Finally, it has been determined that when emotion is not taken into account in decision-making, consequences can be highly disruptive for the person (Bechara, 2004; Rolls, 2014).

Emotional awareness, though, goes a step forward in considering the importance of emotion at the intra-personal level, because it presupposes that learners can benefit from being aware of their own emotions (Lavoué et al., 2020). Whether consciousness is a necessary condition for emotion to *exist* (LeDoux & Brown, 2017; Lieberman, 2019b) is a debate outside the scope of the present contribution (see also next chapter): the use of voluntary self-report requires consciousness for emotional awareness to emerge in the first place and therefore *dodges* the issue. Conscious processing derived from emotional awareness has been related more specifically to three overlapping factors in learning processes (Lavoué et al., 2020): (1) emotion as useful information to direct or redirect cognitive resources; (2) emotion as useful information to assess and guide learners' interest and motivation; and (3) emotion as useful information to regulate one's own behavior, including one's own emotions.

Learners may use their achievement, epistemic and topic emotions as useful cues to evaluate and direct cognitive processes (Pekrun & Perry, 2014). The work of D'Mello and colleagues (2014), for instance, highlights how confusion is a signal of cognitive disequilibrium, which learners are therefore incited to re-balance. The work of Vogl and colleagues (2019) about the role of epistemic emotions suggests that experience of curiosity, pride and shame may inform learners about different reasons for knowledge exploration.

Awareness of one's emotions may prompt learners to assess their motivation to keep engaged or disengage from a learning activity (Linnenbrink-Garcia, Patall, & Pekrun, 2016). For instance, Baker and colleagues' work (2010) draws attention on the importance of recognizing boredom and frustration as warnings of inefficient and potentially misleading efforts.

Finally, emotion may be used as useful information to regulate one's own behavior, including emotion regulation itself (J. J. Gross, 2014; J. J. Gross, 2015). When performed explicitly rather than implicitly (Torre & Lieberman, 2018), emotion regulation requires the person to be aware of the emotion to be regulated (Lavoué et al., 2020). According to Gross, emotion regulation "refers to the processes by which we influence which emotions we have, when we have them, and how we experience and express them" (J. J. Gross, 2002, p. 282). The process model of emotion regulation (J. J. Gross, 2002, 2015) posits that regulation can happen at five successive stages:

1. *situation selection*, which consists in avoiding or experiencing events according to their probability of eliciting unwanted or sought after emotions (*e.g.*, *I won't use this Learning Management System, I get angry every time I use it!*);
2. *situation modification*, consisting in attuning the situation once it has been selected or it is forced upon the person (*e.g., I will use it only to read the forum*);
3. *attentional deployment*, by which some characteristics of the situation are considered more relevant than others (*e.g.*, *The individual resources are good, it is the overall system I don't like*);
4. *cognitive change*, through which the focal points of the situation can be reappraised in an attempt to shift, endure or increase an emotional experience (*e.g.*, *Maybe with practice I will end up to find the system useful*);
5. *response modulation*, which occurs once an emotion has already been elicited and the person attempts, for instance, to suppress its manifestations (*e.g.*, *Ok, stay calm, breath and don't shut the browser!*).

The fourth and fifth passages have been often considered as two alternative strategies for emotion regulation, with potentially different impact on person's cognition and behavior (Bonanno, Papa, Lalande, Westphal, & Coifman, 2004; J. J. Gross, 2002; Richards & Gross, 2000). In some circumstances, cognitive change can be more beneficial, since it reassesses the situation in what shall be more favorable terms for the person, freeing resources that would otherwise be monopolized by the undesired emotion (Richards & Gross, 2000). In other circumstances, a cognitive re-evaluation of the situation may be too demanding, and the person may have more benefits in trying to suppress the response modulation (Bonanno et al., 2004).

Torre and Lieberman (2018), on the other hand, hypothesize that emotion regulation may occur even implicitly through *affect labeling* (Lieberman et al., 2007; Lieberman, 2019a; Lieberman, Inagaki, Tabibnia, & Crockett, 2011), that is, when people assign a word to the their emotional experience (*e.g.* "I feel *angry*"). In their article, the authors provide an overview of research about affect labeling and state that it "has demonstrated a modulation of emotional output effects in the same experiential, autonomic, neural, and behavioral domains as found in other forms of emotion regulation" (Torre & Lieberman, 2018, p. 117). The authors also highlight, though, that the mechanisms by which affect labeling intervene as implicit emotion regulation are still unclear. In this regard, Torre and Lieberman (*ibid.*) propose four possible mechanisms:

1. *distraction*, for resorting to language shift the attention from the situation itself and therefore attenuates full processing of the eliciting event;

2. *self-reflection*, as a means to initiate an introspection process fostering self-distancing from the emotion;
3. *reduction of uncertainty*, which results from categorizing an intense and often nuanced experience using known and community-shared words;
4. *symbolic conversion*, consisting in events assuming symbolic status through the associated word, which may induce more abstract thinking about the eliciting event.

Emotional awareness may therefore be useful either as explicit (J. J. Gross, 2015) or implicit (Torre & Lieberman, 2018) emotion regulation. It may also help learners in evaluating whether a regulatory strategy may be more or less appropriate given the situation at hand (Lavoué et al., 2020).

To sum up, there is a consistent body of research that highlights how emotion play a prominent role at the intra-individual level in various processes, which are also implicated in learning. Learners may therefore benefit from self-emotional awareness as valuable information upon which deciding how to steer their learning processes.

2.2.2 Related Works

Molinari and colleagues (2016) developed a self-reporting system from an experience sampling method (Csikszentmihalyi & Larson, 2014) perspective, the EMORE-L (EMOtion REport for E-Learning). The tool was used in an ecological setting by 16 university students who voluntarily adopted the system during 15 days of distance learning in a blended bachelor program. The system consists in a short online questionnaire that students were reminded to fill once per day through an email that they received at a time previously concerted with the investigators. The EMORE-L consists in nine short questions, which are meant to reduce the amount of time needed to fill-in the questionnaire, organized in 4 parts: a first part about the situation the students were facing, with information about the activity students were conducting; a second part about the cognitive evaluation of the situation on three dimensions (Control, Value, Activation); a third part about the emotional experience, with the possibility of choosing between eight discrete emotions (*pleasure*, *anxiety*, *curiosity*, *boredom*, *engagement*, *confusion*, *surprise*, and *frustration*), for which students could also provide the intensity; and a last part about the social sharing of emotion, with items related to the wish for students to share their emotions with their colleagues, as well as the mutual knowledge of emotions between the student and their colleagues.

Through the 169 questionnaire that were filled throughout the 15 days, Molinari and colleagues (*ibid.*) were able to assess three main exploratory subjects. For each element, the authors used results of the study to discuss consequences on the implementation of emotion self-reporting tool in a context of emotional awareness.

The first subject under scrutiny concerned what emotions were most frequently associated with the different situations a student may encounter during distance learning. In this regard, results highlight four main activities: reading resources, synthesis of the same resources, individual work, and group work. The three most likely emotion to be experience are *pleasure*, *anxiety*, and *surprise*, whereas students made a scant use of the other five emotions proposed by the list. The authors also evinced from their data that the frequency of emotion changed as a function of the activity. On this ground, Molinari and colleagues suggest that self-report tool should take into account the specific activity that they aim to sustain, avoiding thus generic list of emotions.

The second phenomenon explored in the study investigated to what extent students wished to share their emotions with their colleagues, and in what situation. According to their results, Molinari and colleagues posit that students prefer to share their emotions during individual and group works. Furthermore, students were also more inclined to share their emotions when they experienced *anxiety* or *pleasure*. Based on these results, the authors suggest that self-reporting tools shall not be provided at all time, but only when the activity is likely to elicit emotions learners wish to share.

Finally, Molinari and colleagues assessed the contribution of the dimensional approach of the cognitive evaluation items and of the eight discrete emotions in allowing students to express what they felt. The dimensional vs. discrete emotion approach to emotion self-reporting is a longstanding debate (Mortillaro & Mehu, 2015; Scherer, 2005), with both approaches presenting advantages and shortcomings in terms of user experience and quality of data (see below in the chapter). Results observed by Molinari and colleagues, augmented by feedbacks from students regarding the use of the system, confirm this trend. For instance, the authors highlight the scant use of the different discrete emotions, with only three out of eight being adopted frequently. At the same time, the cognitive evaluation provided through the *control*, *value* and *activation* dimensions also presented shortcomings: *pleasure*, *anxiety* and *surprise*, in spite of their diversity, were in fact associated to similar appraisal profiles. On this ground, the authors point out how important it is that the self-reporting system helps students in identifying their emotional experience.

Molinari and colleagues (*ibid*) exploratory work provide useful elements with respect to emotional awareness in a voluntary self-reporting setting. On the technical standpoint, the EMORE-L combines antecedents of emotions, in the form of cognitive evaluation on dimensions, with emotional experience expressed as discrete natural language words, even though the two blocks of items are not intertwined in the user experience. Also, even if relatively short, the use of 9 questions to express one's emotions is more suitable for a *scripting* rather than a moment-to-moment emotional awareness.

At a conceptual level, Molinari and colleagues introduced emotional awareness in individual settings, as a means to provide students with self-awareness of their own emotions. At the same time, the authors measured students' wish to dispose of collective emotional awareness. To this extent, Molinari and colleagues highlight how students wished to have full control over when and which emotional information to disclose. In addition, they also manifested the wish of being aware of their colleagues emotions, that is, bringing emotional awareness at the inter-personal level, which is the subject of the next section.

2.3 Emotional Awareness at the Inter-Personal Level

The second assumption about the instrumentality of emotional awareness posits that learners can benefit from emotional information available at a collective level, either for their own learning processes, that of their partners, or both at the same time.

2.3.1 Theoretical Underpinnings

Whereas the intra-personal function of emotion has received extensive attention in the last few decades, several researchers have more recently advocated the need for investigating the inter-personal function of emotion (A. H. Fischer & Van Kleef, 2010; Parkinson et al., 2005; Rimé, 2005; Van Kleef, 2018). Fisher and Van Kleef (2010), for instance, posit that

> It is an indisputable fact that emotions are mostly reactions to other people, that emotions take place in settings where other people are present, that emotions are expressed towards other people and regulated because of other people. It is hardly possible to imagine the elicitation of anger, shame, sadness, happiness, envy, guilt, contempt, love, or hatred without imagining other people as cause, target, or third-party observer of these emotions. In other words, the social constitution of emotions is beyond doubt.
> – A. H. Fischer & Van Kleef (2010), p. 208.

Keltner and Haidt (1999) identify four level of analysis in which emotion play a social function: (1) the individual level whenever the source of the emotion is of a social nature; (2) the dyadic level, such as in direct dialogue or collaboration; (3) the group-level, in which people share common identities and goals to attain; and (4) the cultural level, where the analysis focus on macro-elements such as history and tradition. An overarching synthesis on the social functions of emotions across all levels is proposed by Fischer and Manstead (2016), who identify two complementary, but distinct, functions performed by emotions: affiliation and distancing. The affiliation function serves to form and maintain positive social relationship with others, whereas the distancing function helps in establishing and maintaining a social position relative to others (*ibid.*).

Other social functions of emotion comprise: the social sharing of emotion (Rimé, 2005, 2009) as a means to initiate and reinforce social-bonding; harnessing emotional information to understand causes and consequences of behaviors in others (Parkinson, 2010; Van Kleef, 2009); the use of emotional information as a reference for normative conduct or for inferring personality traits (Hareli & Hess, 2010; Hareli, Moran-Amir, David, & Hess, 2013); and the manifestation of emotion as a form of involuntary contagion (Barsade, 2002; Barsade & Knight, 2015) or deliberate influence on others (Van Kleef, Heerdink, & Homan, 2017; Van Kleef, Van Doorn, Heerdink, & Koning, 2011). In this regard, Van Kleef (2009, 2010, 2018) proposes an overarching framework, the Emotion As Social Information (EASI) model, which attempts at integrating the many inter-personal functions of emotion in terms of two classes of mutually influential, but conceptually distinct and empirically separable mechanisms:

1. *Emotional expressions trigger affective reactions in observers.* The first mechanism implies that the emotion of a person may be the eliciting stimulus of an emotion in the observer. The affective reaction in the observer may be the same as the one expressed by the person, as in the case in emotional contagion (Barsade, 2002) or empathy (Bloom, 2016; Wondra & Ellsworth, 2015). For instance, boredom may propagate between learners as the result of the first person pausing and puffing. At the same time, the emotion in one person can trigger a different affective reaction

in the observer, but independently of inferential processes about the situation. For instance, the amusement of a colleague during a collaborative task may trigger anger in the observer who wants to stay focused, independently of the reasons why the colleague is amused. Whether the observer affective reaction mirrors or complements that of the person who has expressed it, the reaction will have consequences on the observer's behavior (*e.g.*, deciding to ignore the amused colleague) and/or both the observer and the person who expressed the *original* emotion (*e.g.*, both learners decide to take a break).

2. *Emotional expressions elicit inferential processes in observers.* The second mechanism implies a more complex and deliberate act of information-processing by which the observer attributes to the expressed emotion potential causes of and consequences on behavior. Especially in appraisal theories of emotion (Moors, 2014; Moors et al., 2013; Roseman & Smith, 2001), which will be more thoroughly depicted in the next chapter, the evaluation a person makes of the situation is pivotal in determining what emotion she may feel. By a form of *reverse engineering* (Hareli & Hess, 2010; Scherer & Grandjean, 2007a), the specific emotion a person is feeling may be used to infer how that emotion came to be elicited.

The EASI model (*ibid.*) also specify two preconditions and two moderator factors that determine the inter-personal dynamics of emotion. The preconditions concern (1) the ability of the person experiencing the emotion to *encode* it in a form that can be communicated to the observer, which broadly refers to the concept of *emotion expressivity* (Kring, Smith, & Neale, 1994; Scarantino, 2017); and (2) the observer's ability to *decode* the emotional information and make sense from it using emotional knowledge and understanding, which closely relates to the principle of socio-emotional competences depicted in Section 1.1.3 (Brackett et al., 2019; Cherniss, 2010; Matthews, Zeidner, & Roberts, 2007). The two preconditions reflect the displaying and the monitoring functions of awareness tools depicted in Section 1.3.1 (Buder, 2011; Schmidt, 2002). As stated by Van Kleef (2018):

> No matter how informative emotions may be, and however critical their social-regulatory functions, if emotions are not expressed and/or fail to be perceived by others, their signaling function is evidently lost and social interaction may be jeopardized.
> – Van Kleef (2018), p.52

The EASI model also proposes two classes of moderating factors that may influence to what extent an expressed emotion results in an affective reaction and/or an inferential process by the observer (*ibid.*). For instance, it may be the case that the *confusion* expressed by a colleague may trigger an affective reaction of *anger* in the observer, because this slows down the collaboration. On the other hand, the same expression may also push the observer to understand the causes of the colleague's *confusion*, for then proposing some form of co-regulation. According to the EASI model, the engagement of the observer in deliberate and more cognitively demanding inferential processes may be facilitated by (1) the observer's motivation and capacity to allocate resources to make inferences about causes and consequences on the sender's behavior; and (2) the perceived *appropriateness* of the expressed emotions according to the observer's criteria of evaluation, for instance in relation to social norms (A. H. Fischer & Manstead, 2016; Hareli & Hess, 2010). In other words, the observer may try to understand the colleague's confusion if (a) the observer has some *free* resources to dedicate, which may be difficult if the learning context is demanding, and (b) the observer evaluates *confusion* as an appropriate reaction to the situation, for instance by reckoning that the learning task is difficult.

The EASI theory (*ibid.*) provides an integrative framework of the social-function of emotion, which is founded on the assumption that one's own emotions may guide thoughts, actions and feelings in another person. Using emotion as social information may, for instance, be a first step in the social regulation of emotion, that is, the attempt to voluntary act upon other people's emotion (Netzer, Van Kleef, & Tamir, 2015; Reeck, Ames, & Ochsner, 2016; Zaki & Williams, 2013). According to Reecks and colleagues (2016, p. 48) , "[t]he goal-driven nature of social regulation distinguishes it from related phenomena, such as social sharing, empathy, or emotional contagion, where one person's actions are not strategically directed towards influencing another's emotions". The authors propose a cross-disciplinary model that implements the process model of emotion regulation (J. J. Gross, 2002, 2015), illustrated in the intra-personal section, from an inter-personal perspective. In what they call the Social Regulatory Cycle, depicted in Figure (**ref?**)(fig:tf-social-regulation-model), the authors identify the regulator and the target person upon which the emotion regulation is intended. The regulator must proceed, first, by identifying the target emotion. Second, the regulator must assess whether the identified emotion corresponds to a suitable emotion or not. Third, in case of a discrepancy between the two, the regulator must plan a strategy aiming at producing in the target the suitable emotion. Last, this strategy must be implemented. At this point, the strategy may target a different stage of the process in the target (see the description of the individual model above for more details): the selection or modification of the situation; the orientation of the target's attention; the possibility of changing the way the target appraises the situation; or the modulation of the behavioral, physiological and experiential manifestation of the emotion. The cycle may start again either from an individual standpoint (*e.g.*, the target is dissatisfied with the new emotion induced by the regulator) or inter-individual perspective (*e.g.*, the regulator did not obtain the suitable emotion or identify an even more adapted emotion for the target).

To sum up, there is extensive work in emotion theory advancing reasons why emotions play a pivotal role at the inter-personal level. Learners may therefore benefit from socially conveyed emotional awareness both at the individual and collective level, assuming *appropriate* communicating and inferential processes are at play.

2.3.2 Related Works

Eligio, Ainsworth and Crook (2012) carried out two intertwined experiments aiming at exploring "what collaborators understand about each others emotions and the implications of sharing information about them" (*ibid*, p. 2046). In the first experiment, the authors asked pairs of same-sex, unacquainted participants to play a collaborative game in a co-located environment, where they shared the same computer equipment. At the end of the collaboration, participants were asked to fill in two questionnaires where they had to rate

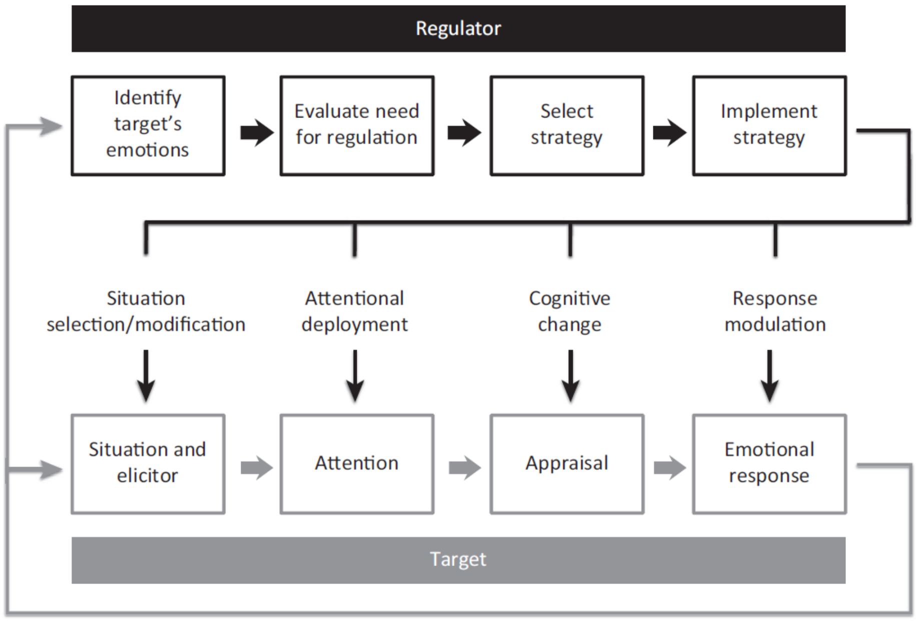


Figure 2.1: The Social Regulatory Cycle at the inter-personal level, fromm the original Figure 2 in Reeck et al. (2016), p. 51.

the intensity of 15 emotions: *happy*, *angry*, *sad*, *fearful*, *angry*, *bored*, *challenged*, *interested*, *hopeful*, *frustrated*, *contempt*, *disgusted*, *surprised*, *proud*, *ashamed* and *guilty*. In the *own* version of the questionnaire, they rated the intensity of each emotion as they have perceived it during the task. In the *partner* version of the questionnaire, participants had to project the intensity by which their partner in the dyad had experienced each of the listed emotion. The administration of the questionnaire was made individually, so that participants could not discuss the matter with their partner, and up until that moment, participants were not informed of the rating, so that they had no particular reason to pay attention neither to their own, nor to their partners' emotions. By comparing the responses to the two questionnaires, the authors concluded that, despite collaborating side-by-side, participants had little understanding of their partner's emotions. On the contrary, consistently with previous findings in the literature (Krueger & Clement, 1994; Toma, Yzerbyt, & Corneille, 2010), participants tended to project their own emotional experience onto their partners. Eligio and colleagues (*ibid*) provides two possible explanations for these results. On the one hand, participants could simply not care about the emotions of their partner. On the other hand, participants could genuinely care about them, but lack the means to focus on them, for instance because the task was too demanding, and they decided to prioritize other mental states perceived as *more* instrumental to the task at hand. In both cases, the attribution of their own emotional experiences to that of the partner is caused by a lack of information available, but the reasons for this shortcoming are not the same depending on the intentions.

In the second experiment, Eligio and colleagues (*ibid*) therefore decided to intervene by providing (or not) participants with explicit awareness of their partner's emotions using a *scripting* strategy, that is, by interrupting the collaboration at specific moments for participants to express their emotions and projecting that of their partner. After filling the *own* and *partner* questionnaire, if participants disposed of awareness, they could look at the emotions of their partner's before resuming the task, otherwise they just continued with the collaboration. Furthermore, dyads were also assigned either in a co-located or a remote collaborative setting, resulting in a 2x2 factorial design with awareness vs. no-awareness, and co-located vs. remote conditions. Using experimental settings similar to the first experiment (but with a slightly different collaborative task and with only women as participants), the authors report evidence suggesting that participants benefited of emotional awareness in terms of performance both in the co-located and remote conditions. Furthermore, participants in the remote condition were more accurate in understanding emotions of their partner and also experienced more *positive affect*, computed by averaging the intensity of *happy*, *interested*, *hopeful*, *excited* and *challenged*.

From the two experiments, Eligio and colleagues (*ibid.*) concluded that participants do not have an accurate understanding of the emotion of their partner if their attention is not explicitly driven to it. Providing emotional awareness seems thus a promising way to increase mutual understanding, since participants with emotional awareness showed higher accuracy in estimating their partner's emotions, as well as better performance to the task at hand.

Following Eligio and colleagues (*ibid.*) findings, Molinari and colleagues (Gaëlle Molinari, Chanel, et al., 2013) conducted a study in which 30 dyads of same-sex participants (16 dyads of women, 14 dyads of men) performed a collaborative task in remote conditions, with the possibility of audio but not video connection. The aim of the collaborative task was to conceive a slogan against violence in schools using an argument graphic tool (Lund, Molinari, Séjourné, & Baker, 2007). The task therefore implied some form of negotiation for deciding the final outcome of the collaboration, which was intended to create socio-cognitive tensions (Andriessen et al., 2011) in participants as a means to elicit emotions. Half of the dyads were randomly assigned to a control condition, in which participants did not dispose of emotional awareness. The other half of the dyads were provided with emotional awareness through the use of an EAT persistently available on screen alongside the collaborative tool to build the slogan. (The tool will be described in more detail below, see Figure 2.4.) Contrary to Eligio and colleagues (*op. cit.*), in which participants shared emotions at specific moments during the task following a *scripting* strategy, in this case the setting was congruent with an *awareness* strategy, since participants could express and have access to their partner's emotions at any time during the task. A message also showed up after 5 minutes from the last expressed emotion to remind participants to express how they were feeling.

At the end of the collaborative task, participants individually filled-in a questionnaire aimed at gathering information about (1) the kind and intensity of the emotions of the participant at the end of the task, as well as the emotions the participant attributed to the partner, particularly with respect to the 20 discrete emotions available as buttons on the EAT; and (2) the quality of the perceived interaction based on the frequency by which the participant and the partner provided/imposed their own points of view, defended and argued their ideas, understood their partner's points of view, built up on their partner's ideas, as well as managed emotion during interaction. As for Eligio and colleagues (*op. cit.*), also Molinari and colleagues (*ibid.*) report that their result support the hypothesis of beneficial effect from the presence of an EAT, but only in dyads of women and with the presence of mixed results with respect to their initial hypothesis.

The tests upon which Eligio and colleagues (*op. cit.*) and Molinari and colleagues' (*ibid.*) results are based, though, do not consider the hierarchical structure of the dyads, which may inflate the rate of type I error since the non-independence of observations is not taken into account by the ordinary least square models they have adopted in their analyses (Brown, 2021; Singmann & Kellen, 2020; West, Welch, Gałecki, & Gillespie, 2015). As a consequence, the evidence provided by the two studies should be taken with caution. The experimental settings, on the other hand, are interesting and directly compare a *scripting* and an *awareness* perspective, with the latter presenting the advantage of persistent, moment-to-moment emotional awareness. The two studies also highlight the many methodological challenges in determining the effect of an EAT in computer-mediated learning environments, which concern how emotional awareness is provided, but also how the potential benefit are measured and analyzed.

A different approach was taken by Avry and colleagues (2020), who, rather than providing awareness through discrete emotions, took a dimensional approach adopting Pekrun's Control-Value theory of achievement emotions (Pekrun, 2006) introduced in Section 1.1.1. In their study, the authors asked 28 dyads of same-sex participants to play a remote collaborative game. Beside the game window, participants disposed of a smaller window on the screen which reported two feedbacks: (1) a feedback on how well the dyds was mastering the game, representing the *Control* dimension; and (2) how well they were performing in a standing compared to other dyads, representing the *Value* dimension. The two feedbacks were manipulated by the authors as to create a 2x2 factorial design combining high vs. low *Control* on the one hand, and high vs. low *Value* on the other. After the collaborative task, participants filled in a questionnaire, on which they evaluated the overall collaboration in terms of (a) affective dimensions (Valence, Dominance, and Activation) and 16 discrete achievement emotions, and (b) six computer-supported collaborative exchanges (sustaining mutual understanding, information pooling, transactivity, reaching consensus, task management, and time management). For both category of measures, the rating was performed for the participant herself, as well as what the participant thought her partner would have experienced. The results obtained by Avry and colleagues (*ibid.*), who took into account the hierarchical structure of dyds by checking the intraclass correlation of dyads, corroborate an effect of the manipulation of the Control-Value appraisals on both category of measures (affective and socio-cognitive). These results are particularly interesting considering that in this study a form of emotional awareness was (1) manipulated, and (2) conveyed through a feedback aimed at eliciting a certain kind of appraisal of the situation, which influenced the kind of discrete achievement emotions experienced.

The three studies outlined in this section provide very different approaches to the study of emotional awareness in computer-mediated learning environments. Since this is a recent and cross-disciplinary field of inquiry, a fragmented stage of research is inevitable (Fiedler, 2004). On the other hand, efforts should also be dedicated to the possibility of direct comparison and incremental building of knowledge. In this regard, the adoption of a prototypical EAT could be beneficial, for similarities and differences in conveying emotional awareness can be more easily defined through objective parameters.

2.4 Emotion in Computer-Mediated Learning Environments

The third assumption underlying emotion awareness tools concerns the possibility to create emotional awareness in computer-mediated learning environments.

2.4.1 Theoretical Underpinnings

For emotional awareness to emerge and be available, it is necessary that emotion may be fruitfully *encoded* and *decoded* in a computer-mediated learning environment. In this regard, the computer-mediated environment is somehow ambivalent with respect to affect-related phenomena. On the one hand, the combined presence of a learning task and a technological device captures learner's attention and thus, even in the case of co-located interaction or seamless audio/video connection in remote settings, diminishes the possibility to pay attention to efferent cues of affective experiences – such as facial expressions, vocal prosody or body posture (Bänziger et al., 2009) – that are more readily available in face-to-face interaction (Baltes et al., 2002; Lund et al., 2007). On the other hand, the same presence of a technological device provides alternative or complementary means to create emotional awareness and even extend it over time (Derick et al., 2017; Derks et al., 2008; Glikson et al., 2018; Hegarty, 2011; Leony et al., 2013).

The presence of emotion in computer-mediated learning environments even without seamless audio/video connection can be sustained by a corollary to the EASI model, illustrated in the previous section, which Van Kleef (2017) identifies as the *functional equivalence hypothesis*. The hypothesis posits that the role emotion play at the inter-personal level is equivalent regardless of the specific way it is communicated, as long as the emotional information passes from the sender to the receiver. In the words of the author:

> If one accepts the notion that emotional expressions can influence social interactions by providing information about what is on the expresser's mind, it follows that emotions can have such effects regardless of how they are expressed, as long as the expressions convey the relevant information. Consequently, EASI theory posits that expressions of the same emotion that are emitted via different expressive modalities (i.e., in the face, through the voice, by means of bodily postures, with words, or via symbols such as emoticons) have comparable effects, provided that the emotional expressions can be perceived by others
> — Van Kleef (2017), p. 213

On the other hand, the fact that various forms of emotional expression are equivalent with respect to their social function does not mean that each way of expressing an emotion is equivalent (Van Kleef, 2018). A colleague's anger expressed by shouting on your face how badly your part of a collaborative task has been handled has not the same intensity of *I am very angry with you* written in an email. A diminished *veracity*, though, may be compensated by more favorable conditions to engage in inferential processes rather than affective reactions as described by the EASI model (*ibid.*).

An EAT whose aim is to make learners' aware of their own and/or their colleague's emotions must therefore face the challenges of *encoding* and *decoding* emotional information, so that it maximized the preservation of its *functional* meaning.

In this regard, the *encoding* of emotional information has received so far greater attention than *decoding*, since it is tightly related to the action of detecting, representing or measuring emotions (da Silva et al., 2020; Fuentes et al., 2017; Mauss & Robinson, 2009; Mortillaro & Mehu, 2015). In their review of technologies for emotion-enhanced interaction already cited in Section 1.2, Cernea and Kern (2015) distinguishes between three types of techniques commonly adopted to estimate emotions: (1) perception-based estimation, derived from efferent manifestations of emotion such as facial expressions, vocal prosody or body posture (Bänziger et al., 2009; Martinez et al., 2016); (2) physiological estimation, based on the detection of physiological patterns such as heart rate, blood pressure, or skin conductance (Ragot, Martin, Em, Pallamin, & Diverrez, 2018; Shu et al., 2018); and (3) subjective feelings, based on the person's self-report of her own emotional experience (da Silva et al., 2020; Lavoué et al., 2017; Ritchie et al., 2016). As stated in the introduction, the thesis focuses on this last category. In fact, physiological estimation requires dedicated hardware and software, which would be difficult to provide at scale. Perception-based estimation may use more widely available hardware and software, for instance through a webcam or keyboard stroke, but the accuracy and usefulness of this kind of measure is still lively debated in the literature (Bahreini, Nadolski, & Westera, 2016; Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019; M. Nazmul Haque Nahin, Alam, Mahmud, & Hasan, 2014). Barrett and colleagues (2019), for instance, argue that automatic recognition from facial expression are still limited in reliability, lack in specificity, and does not take sufficiently into account effects of context and culture. Voluntary self-report is therefore retained as the most parsimonious, portable and reliable way to provide emotional awareness both at the intra-personal and inter-personal levels, especially when the aim is to stimulate voluntary emotional introspection and/or inferential processes (Boehner et al., 2007; Van Kleef, 2018).

Emotional self-report is tightly related to the underlying concept of *what an emotion is*, which will be discussed in Chapter 3. From a technical standpoint, emotion self-report is traditionally characterized by the dimensional or the discrete emotion approaches, as well as by a combination of the two (Bradley & Lang, 1994; Cowen & Keltner, 2017; Mortillaro & Mehu, 2015; Scherer, 2005).

In a dimensional approach, an emotion is conceptualized as a rating on a number of continuous criteria, among which the most frequently adopted are *Valence* (also called *pleasure*) and *Arousal* (also called *activation*) (*e.g.*, James A. Russell, 1980; Stanley & Meyer, 2009). Different or additional dimensions are nevertheless also possible. It is the case, for instance, of the Self-Assessment Manikin (Bradley & Lang, 1994) or the AffectButton (Broekens & Brinkman, 2013), which both use the *Valence*, *Arousal*, and *Dominance* dimensions, even if in different formats. The Self-Assessment Manikin (Bradley & Lang, 1994) uses three rows of figures (*i.e.*, the *manikin*), where each row represents one of the three dimensions: 5 figures, progressively modified in some features, represent the increasing or decreasing value on the specific dimension. The AffectButton (Broekens & Brinkman, 2013), on the other hand, uses a single iconic facial expression that changes according to the user's coordinates of the mouse on the surface of the icon: the horizontal movement determines the pleasure dimensions; the vertical movement the dominance dimension; and the arousal dimension is calculated according to the distance of the mouse from the central point of the image.

In the discrete approach, emotion is conceptualized as a phenomenon with distinctive features compared to other emotions, which may consist in different facial expressions when emotion is represented graphically, or semantic meaning when emotion is represented by natural language words or idioms (Desmet, 2003; Fontaine et al., 2013; Ruiz et al., 2016; Torre & Lieberman, 2018). The number, kind and organization of the representations varies depending on several aspects. In the case of verbal representations that adopts natural language nouns or adjectives, for example, the list may be determined according to emotion theories – as in the case of basic emotion theory (Ekman, 1992) – or determined empirically according to previous (or pilot) studies, often with the aim of retaining a list of the most frequently experienced or expressed items. For instance, Molinari and colleagues (2013) implemented in a self-report interface 20 buttons, 10 labeled with *negative* and 10 with *positive* emotion adjectives retrieved among the most frequently expressed during a computer-mediated collaborative task in a pilot study. As alredy mentioned in Section 1.1.2, Ruiz et al. (2016) adopted a similar approach with 12 discrete emotions. When emotion is conceptualized graphically, representations usually attempt at maintaining some degree of analogy with *reality*, for instance with respect to facial expressions (Desmet, 2003; Laurans & Desmet, 2012). In the discrete emotion approach, respondents can for instance: (1) choose the item from a predefined list that best matches the emotional experience; (2) rate their agreement on a scale for each item in a predefined list as to the extent they have experience that particular discrete emotion; and (3) rate the intensity by which each discrete emotion has been experienced (Scherer, Shuman, Fontaine, & Soriano, 2013).

The dimensional and discrete approaches can be combined by providing discrete emotions a precise collocation on the dimensions. The result is often referred to as an *affective space* (Gillioz et al., 2016; Scherer et al., 2006; Shuman & Scherer, 2014). Recent work on the meaning of emotional terms seems to suggest that four dimensions are needed to account for emotion differentiation : *Valence*, *Power*, *Arousal*, and *Novelty* (Fontaine & Scherer, 2013; Fontaine et al., 2007; Gillioz et al., 2016). Gillioz and colleagues (Gillioz et al., 2016), for instance, empirically mapped through a principal component analysis 80 french emotions terms on the four-dimensional affective space. (More on this in Section 3.4.4.)

Emotional *decoding* in a computer-mediated environment is tightly linked on how emotion is *encoded* and has received so far limited attention (Berset, 2018; Derick et al., 2017; Leony et al., 2013). Furthermore, visualization of emotion in this context is usually derived from affective information available in students' productions or communication exchanges, for instance through sentiment analysis rather than a dedicated tool (Leony et al., 2013). In the absence of specific representations for emotional awareness, more general guidelines about the visuo-spatial representation of data should be applied (Hegarty, 2011; Hehman & Xie, 2021). Emotional *decoding* is also influenced by the objective of the visualization.

Learners' emotion may be represented individually or collectively, as a single unit in time or grouped through short or long time spans. They may also be partially or fully available to all other students or only to a selection of colleagues (e.g. in a group work). In other words, if the aim of emotional *encoding* is rather straightforward, emotional *decoding* depends on a larger number of factors.

2.4.2 Related Works

Henritius and collegues conducted a systematic review of empirical research conducted on universities' students emotions in virtual learning based on 91 articles published between 2002 and 2017 in four journals (Henritius, Löfström, & Hannula, 2019). Among the conclusions stated from this review, the authors state that "on a more critical note, we observed that only a few of the studies actually pay attention to the fluctuation of emotions in the context and in the flow of events" (*ibid.*, p. 97). According to the authors, most studies analyzed emotions retrospectively and conceptualize them as traits rather than transitory states or processes (*ibid.*).

The lack of studies investigating affect or emotion in a dynamic perspective is also due to the fact that there is scant work that has been dedicated to instruments that combine both emotion *encoding* and *decoding* in a computer-mediated learning environment, especially in real-time (Ez-zaouia et al., 2020; Lavoué et al., 2017). As pointed out by Lavoué and colleagues (Lavoué et al., 2017), the few attempts that have been made are mostly ad-hoc solutions that do not aim at a general application. As a result, in learning settings, emotion self-report is often provided through questionnaires (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011; Pekrun, Vogl, Muis, & Sinatra, 2016) or adaptation of experience sampling methods (Csikszentmihalyi & Larson, 2014; Gaëlle Molinari et al., 2016; Scollon, Kim-Prieto, & Scollon, 2003).

Emotion representation, on the other hand, is even less frequent and therefore less developed, especially in real-time (Berset, 2018; Derick et al., 2017; Ez-zaouia et al., 2020; Leony et al., 2013). Ez-zaouia and colleagues (2020), for instance, developed EMODASH, a dashboard for visualizing retrospective emotions for tutors in an online learning environment. Leony et al. (2013) and Derick et al. (2017) also propose dashboard inserted into a computer-mediated learning environment, but limited to a handful of emotions (e.g. *frustrated*, *confused*, *bored*, *happy*, and *motivated),* which are automatically derived from learners' activities. Leony et al. (2013), in particular, propose a series of visualizations organized as time-based*,* context-based*,* visualizations of change in emotions*,* and visualization of accumulated information.

To the best of my knowledge, there is only a handful of tools reported in the literature that come close to the purpose of an EAT as considered in the present contribution. I will focus on three of them in the reminder of this section.

Feidakis and colleagues (Feidakis et al., 2014; Feidakis, Daradoumis, Caballé, Conesa, & Gañán, 2013) developed the *emot-control* emotion aware system, depicted as a pop-up window in Figure 2.2. The tool combines the dimensional and discrete approaches by placing 12 icons of facial expressions *(inspired, excited, interested, relaxed, curious, confused, anxious, indifferent, bored, tired, angry, desperate*) in Russel's (1980) *Valence* x

Arousal/Activation circumplex, with a neutral face in the middle. The tool also presents a text input, which originally shows the noun associated to the icon on which users have clicked, but that can also be switched to another semantic word/expression. Users can also provide more context to their feeling by typing a description through a second text-area. 5 other icons on the right-hand of the interface represent each a different mood (*sad*, *unhappy*, *neutral*, *happy*, and *very happy*). Users can access their own last affective episode reported directly on the interface, whereas they have to click on a button to open a new window to access their colleagues' emotions. The window shows a vertical timeline of affective episodes organized by group members, with the date, the time, the emotion (combination of icon and noun/expression), and the mood. In a study, Feidakis and colleagues (Feidakis et al., 2013) adapted the tool to the Moodle learning environment and, in accordance to the Affect-Aware perspective (see previous chapter), endowed the environment also with an Affective Virtual Assistant, which responded to learners affective episodes. Congruently with the approach advocated in this contribution, learners where free to use the tool as they wished, and they could therefore dispose of emotional awareness at any time in the learning environments. In another, similar, study (Feidakis et al., 2014), the authors administered the System Usability Score Brooke (1996) to measure the perceived usability of the tool. With $N = 29$ participants, they obtained a rating of $M = 67.91$ (no SD provided) out of 100. (More about this in Chapter ??.)

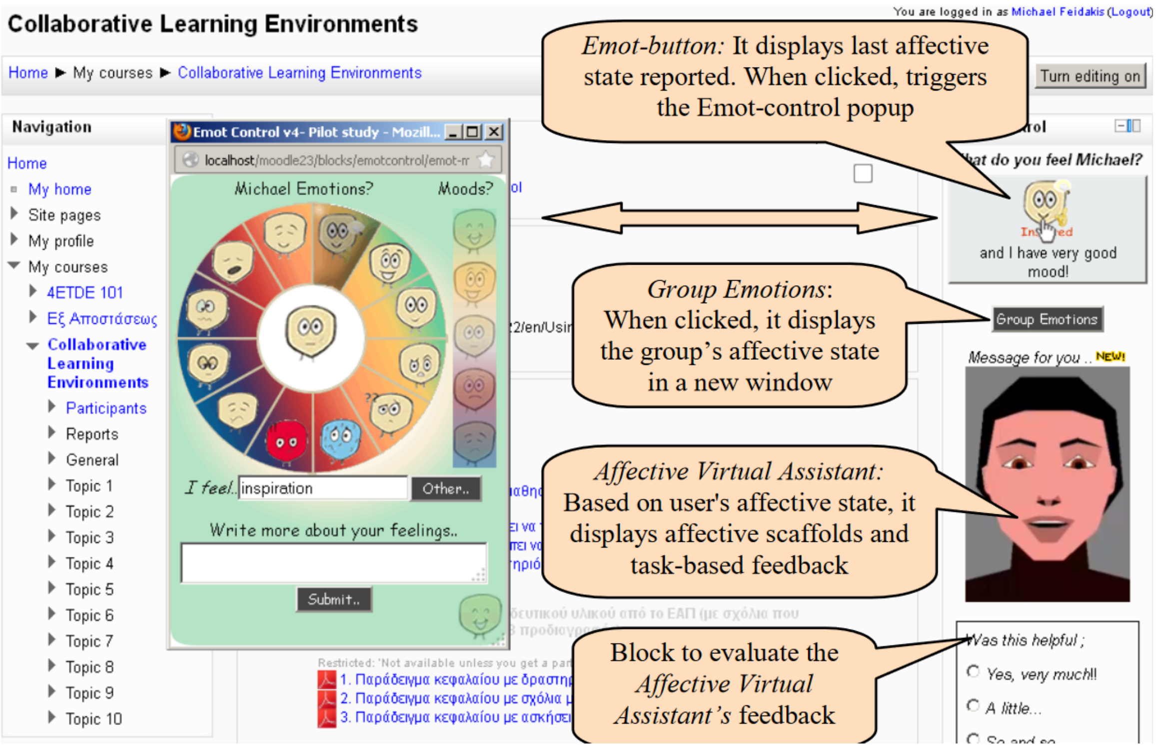


Figure 2.2: Image of the overall interface in which the *emot-control* is inserted. Retrieved from Feidakis et al. (2013), Figure 7 in the original article. p. 1653.

The tool provided by Feidakis and colleagues comply with all the three main features of an EAT as advocated by the present contribution: it is based on voluntary-self report, it incorporates an emotional structure based on Russel's (1980) circumplex, and allows moment-to-moment emotional awareness. On the other hand, the tool separates the expressing-displayingand the perceiving-monitoringfunction of an awareness tool, which are provided in two separate windows. Furthermore, the use of icons representing synthetic facial expressions present some limitations. From a spatial perspective, augmenting the number of options would reduce the size of the icons and make identification more difficult. Depicting emotions with icons is also more prone to misunderstanding or difficulty to provide a sufficient number of images, and different enough to discriminate between similar feelings. For instance the emotion placed from noon to 1 (*inspired*) in the circumplex is very similar to that placed from 2 to 3 o'clock (*interested*). It is useful to note in this regard that a previous version of the tool combined the icon and the label, for the label being dropped in the latest version to overcome language barriers (Feidakis et al., 2013). Finally, the tool combine emotions and moods on the same observation, which may be problematic given a growing consensus in considering the two as distinct phenomena: we may have moods and emotions at the same time, but also only moods or only emotions (Linnenbrink-Garcia et al., 2016; Scherer, 2005, 2022).

The Mood Meter Mobile Application[1], depicted in Figure 2.3, is a digital extension of the physical Mood Meter dashboard of the RULER approach (Brackett et al., 2019; Hoffmann et al., 2020; Nathanson et al., 2016) introduced in Section 1.1.3. The physical Mood Meter consists in four colored quadrants organized according to two axes: *Valence* and *Activation/Arousa*l. Learners can therefore project themselves in one of the four quadrants according to (1) the extent by which they rate their emotional exprience as pleasant or unpleasant, and (2) how high or low is their energy. The top-right yellow quadrant corresponds to emotions which are pleasant and high on energy, such as *excitement*, *joy* and *elation*. The bottom-right green quadrant corresponds to emotions which are pleasant and low on energy, as *tranquility*, *serenity*, and *satisfaction*. The bottom-left blue quadrant corresponds to emotions which are unpleasant and low in energy, as *boredom*, *sadness* and *despair*. And finally, in the top-left red quadrant corresponds to emotions which are unpleasant and high in energy, such as *anger*, *frustration* or *anxiety*. In the digital app, the overall dashboard is divided in 100 points (25 per quadrant). Users can tap one of the point in the quadrant and obtain 9 emotion words, that is the word that is associated with the very point they have clicked, plus 8 alternatives, which corresponds to the 8 adjacent points on the dashboard. Once identified the emotion term that corresponds to how the learner is feeling, the person can: (1) describe the feeling by typing in more information; and (2) select a strategy between quotes, images or practical tips that are meant to help the person shift to another feeling if desired. Further features of the app include the possibility to track the feelings entered in the app, for instance with an overview of the percentage per quadrant; set a reminder to use the app at specific periods; and finally share their feelings via Facebook or Twitter. As the *emot-control*, also the Mood Meter app comply

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

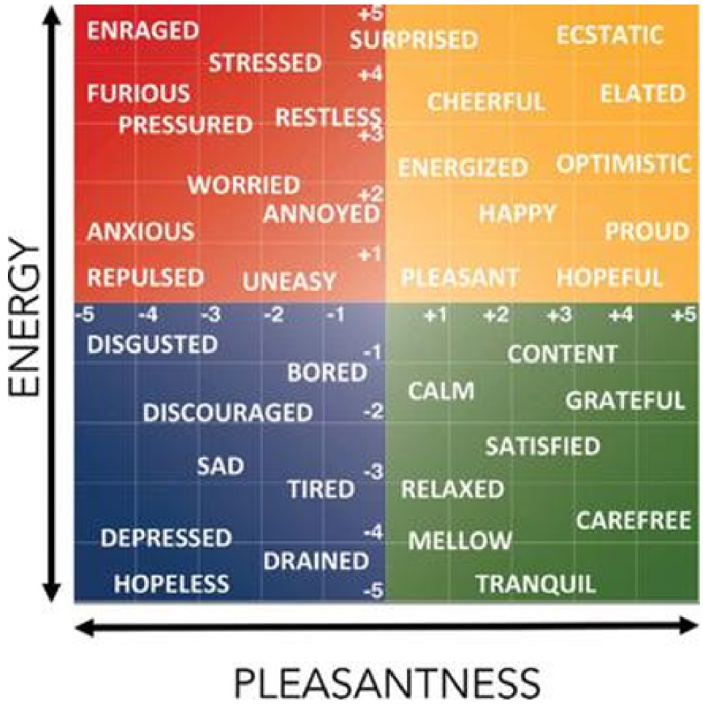


Figure 2.3: Image of the four quadrants of the Mood Meter, retrieved from Brackett et al. (2019), Figure 1 in the original article. For the Mood Meter app, see the official site in the footnote.

with the three main features of an EAT advocated in the present contribution, for they both are self-report tools, based on the same underlying emotion structure, and can be applied in a moment-to-moment perspective. On the other hand, the Mood Meter app seems to be more oriented towards intra-individual emotional awareness: the possibility to share the emotion via a social platform is a very limiting feature to provide group-specific and ongoing inter-personal awareness. Furthermore, to the best of my knowledge, the app is provided only in a mobile version, which would force learners to switch back-and-forth from the computer-mediated learning environment to their phone.

A third and final tool cited in this overview is the one adopted in the contribution of Gaëlle Molinari, Chanel, et al. (2013), already discussed in the related works of Section 2.3 about inter-personal awareness. As a reminder, half of the dyads in a computer-mediated collaborative task disposed of a persistent EAT on the right-hand side of the screen. As shown in Figure 2.4, the EAT is vertically divided in two areas. On top, the monitoring-perceiving area consists in the last three discrete emotions expressed by the participant (green boxes) and the partner (blue boxes). The box on top of each pile, representing the very last emotion expressed, is highlighted with a paler color. The lighter green box is also editable by the participant, who can type-in an emotion not available through the buttons in the lower part of the screen. These buttons represent the displaying-expressing function and are organized in two columns: the right one with 10 *positive* emotions, and the left side with 10 *negative* emotions. The list of discrete emotions were defined mixing a previous study in the field (D'Mello & Graesser, 2012) and 2 pre-experiments that aimed to identify the most frequent and intense emotions felt during a situation of collaboration, real or imagined.

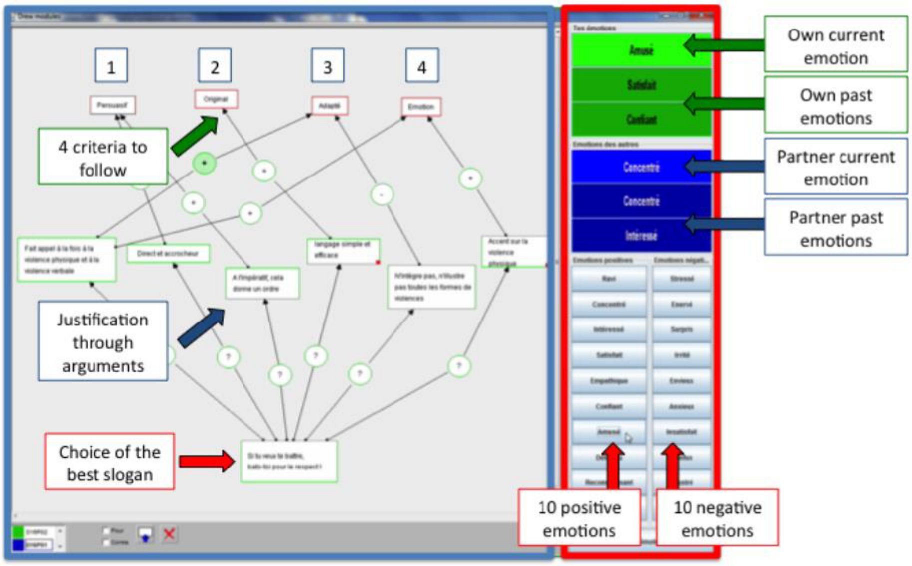


Figure 2.4: The argument graphic tool and the EAT adopted in Gaëlle Molinari, Chanel, et al. (2013), from Figure 1 in the original article.

The EAT adopted by Molinari and colleagues (*ibid.*) is based on self-report and provide moment-to-moment emotional awareness, being persistently on screen. On the other hand, though, it adopts a purely discrete approach to emotion, without implementing any explicit emotion structure into the tool beside the dichotomous organization of emotion terms according to the *positive* vs. *negative* distinction discussed in Section 1.1.2 (Colombetti, 2009; Erbas, Ceulemans, Koval, & Kuppens, 2015; Shuman et al., 2013). In this regard, as a follow-up of the experiment, the authors reckoned that the tool could be improved on a number of points, which became the input for my Master thesis (Fritz, 2015) detailed in Chapter 4.

The three EATs illustrated in this section highlight some of the many choices in *encoding* and *decoding* emotional information that can be made in providing emotional awareness through a dedicated self-report tool. These choices can influence emotional awareness both at the intra- and inter-personal levels in ways that are put into perspective in the next section, which provides an abstract model of the functions of an EAT.

2.5 Abstract Model of the Functions of an Emotion Awareness Tool

The information illustrated above, as well as in the previous chapters, can be integrated in an abstract model that depicts the mechanisms allegedly carried out by an EAT based on the three underlying assumptions, namely that emotional awareness is useful at the individual level, at the collective level, and that it can be fruitfully implemented in a computer-mediated learning environment. The proposed abstract model, though, does not have the pretension to be exhaustive. On the contrary, it aims at providing an organized overview of some of the many different affective, learning or human-computer interaction processes that are (or may be) assumed in the instrumentality of an EAT, as well as how design choices can sustain these processes. The model takes into account elements that are specific or particularly important from the perspective of an Emotion Awareness Tool, assuming that the general functions of an awareness tool illustrated in Section 1.3 are also at stake.

The model, depicted in Figure (**ref?**)(fig:thesis-scm-model), take the perspective of a learner that disposes of an EAT in its computer-mediated learning environment and comprises boxes for four main conceptual elements:

- **Learning activity**, which may broadly refer to any computer-mediated environment implementing an instructional design;
- **Intra-personal emotion**, representing a single event or a set of events corresponding to the learners' expressed-displayed emotions;
- **Inter-personal emotion**, representing a single event or set of events corresponding to emotions expressed-displayed by other learners sharing the same environment and that the learner herself can perceive-monitor;
- An overarching ***meaning-making*** **process**, which encompasses learner's effort to extrapolate instrumental information from the emotional information available.

The boxes are connected with directional arrows, numbered from 1 to 7, representing a series of processes that are (or may be) implicated in each passage. The numbers are used for identifying the passages, but do not imply a fixed order, even though some processes may be built upon previous passages. This section describes each passage in more detail.

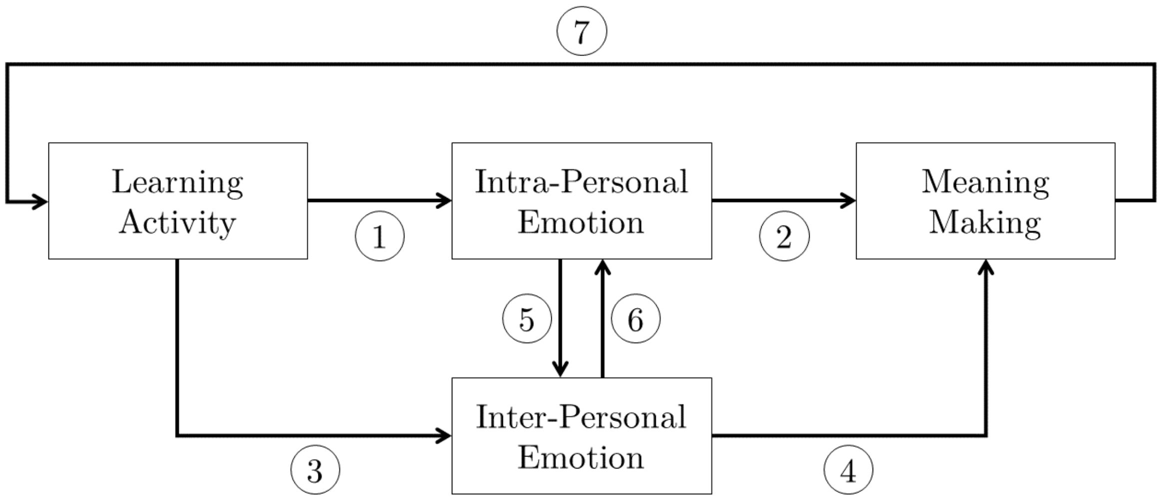


Figure 2.5: Abstract model of the functions of an EAT from the perspective of a single learner disposing of it in a computer-mediated learning environment. Numbers pinpoint passages or processes that may influence whether and how an EAT sustains the function at stake.

2.5.1 From the Learning Activity to Intra-Personal Emotion

Passage #1 goes from the learning activity to an intra-personal emotion, implicating that the learning activity elicits emotions in oneself. An EAT may intervene in this passage at least in two ways, labelled here as emotion alertness and emotion conceptualization for illustration purposes only.

Emotion alertness broadly refers to the fact that the presence of the EAT represents a general reminder about paying attention to emotion during the learning activity, and therefore may increase the *sensitivity* about emotional experience. In other words, the presence of the EAT may push learners to bestow to emotional self-awareness more attention that they would normally do without the presence of the EAT (Brackett et al., 2019; Lavoué et al., 2020; Gaëlle Molinari, Chanel, et al., 2013). Emotion alertness may be induced to different degrees, depending on, for instance, whether the EAT is always available or not, and whether explicit prompts are frequently sent to learners (Csikszentmihalyi & Larson, 2014; Shiffman, Stone, & Hufford, 2008).

Emotion alertness is not necessarily and automatically linked to the concrete action of encoding an emotion into the system. In fact, even in the case of an increased alertness, learners could genuinely not feel any emotion, or decide not to express their emotion through the EAT either for relatively stable *dispositional* characteristics (Kring et al., 1994; Scherer, 2021a) or for contingent reasons (*e.g.*, bestow priority to the learning task at hand).

Emotion conceptualization, on the other hand, broadly refers to the way learners self-report their emotions, which may vary greatly according to a number of theoretical and technical aspects, such as the dimensional, discrete, or combined approach to emotion measurement/reporting. Depending on how emotion is conceptualized and, by extension, planned for self-report, the intra-personal awareness of emotion may be more or less guided, inducing learners to pay attention to some elements of their emotional experience that they will not necessarily consider in the same way, or with the same weight, depending on the specifics of the EAT at hand.

At the same time, emotion self-report also depends on the learners' ability to identify their current emotional episodes and provide an accurate representation of them (Erbas et al., 2015; Scherer, 2009b; Siemer, Mauss, & Gross, 2007; C. Smith & Ellsworth, 1985). Erbas and colleagues (2015), for instance, found evidence that people who differentiate only between a limited set of discrete emotions tend to use *Valence* alone as a discriminating criterion. On the contrary, people discriminating between a larger panel of discrete emotions focus more thoroughly on different features of the situation at hand (see also Chapter 3). An EAT may therefore impinge learners to reflect on a greater variety of potential emotional experiences rather than reinforcing the *positive/pleasant* vs. *negative/unpleasant* loop.

Finally, the EAT may influence learners to focus only on the emotional experience *per se*, or also provide contextual information about the situation or potential antecedents of that experience at various level of details (Feidakis et al., 2014; Lavoué et al., 2020). The richer the details and the reflection to account for them, though, the greater the effort for learners to concentrate and report about the affective component.

To sum up, the presence and features of an EAT may not only contribute to determine whether learners will pay attention to their emotional experience, but also how they will be guided to think about it, planning the way to the next passage.

2.5.2 From Intra-Personal Emotion to Meaning-Making

Once an emotion or a set of emotions are self-reported, it is assumed that the learner may extrapolate meaning from the very emotional information she has provided. In this regard, information may be punctual, for example the last emotion or set of emotions displayed concurrently. Conversely, the system may also implement some form of data aggregation and persistence, which allows to observe the accumulation and the evolution of the emotional experience over a varying time span (Berset, 2018; Leony et al., 2013). Depending on the way emotion information is graphically represented, emotional meaning-making can be guided by the system, which may privilege some form of representation (*e.g.*, changes over time) over another (*e.g.*, cumulative frequency of the same emotion expressed).

Meaning-making is also tightly related to the way emotion is expressed through the system (passage #1), not only from a technical standpoint – since the information available is determined by what information is inserted into the system – but also conceptually. If emotion is considered from a dimensional standpoint, learners will extrapolate meaning from the criteria through which dimensions are rated. On the contrary, in a discrete emotion approach, learners extrapolate meaning from an *information unity* that is supposed to provide unique meaning compared to other possible choices. If the system combines dimensional and discrete approach, learners can extrapolate meaning from both. Moreover, if contextual information is provided, learners can situate emotion more easily; otherwise, inferences on causes and consequences of behavior must be drawn from representations of emotion alone (Van Kleef, 2018).

The form through which emotional information is rendered on screen also influences the cognitive effort necessary to process it (Hegarty, 2011). According to complexity and richness of the graphical rendering, processing can span on a continuum with *almost effortless perception* on the one side, and *deliberate cognitive effort* on the other.

To sum up, the design of the EAT determines the way by which the learner extrapolates meaning from her own emotional information and therefore influence to what extent the learner has interest in monitoring what she has inserted into the system.

2.5.3 From the Learning Activity to Inter-Personal Emotion

Passage #3 determines how the learner becomes aware of emotional information provided by others. It therefore starts a complementary path with respect to passages #1 and #2, but from an inter-personal perspective. This path may in fact be *closed* when the EAT is exclusively concerned with emotion self-awareness (*e.g.*, Lavoué et al., 2020), whereas it becomes an integral part when the EAT is inspired from a Group Awareness Tool perspective (*e.g.*, Avry, Molinari, et al., 2020; Eligio et al., 2012; Feidakis et al., 2014; Gaëlle Molinari, Chanel, et al., 2013).

As for the intra-personal passage #1, the presence of the EAT can have, in the first place, an alertness effect. The tool may prompt the learner to pay more attention to her colleague(s) emotions of what she would have done without the presence of an EAT, as suggested by the results of Eligio et al. (2012) and Gaëlle Molinari, Chanel, et al. (2013) outlined in the empirical works of Section 2.3. In this regard, the design of the EAT may determine to what extent emotion information about others is readily available during the learning activity, as with the EAT adopted my Molinari and colleagues (*ibid.*), or must be voluntarily sought after in another screen, as in the case of the *emot-control* in Feidakis and colleagues (Feidakis et al., 2014; Feidakis et al., 2013). The closer the emotional information is to the task, the greater may be the chances to frequently notice and pay attention to it. At the same time, considering that affective stimuli often are bestowed precedence (Brosch et al., 2010; Pool, Brosch, Delplanque, & Sander, 2015) in cognitive processing, the ongoing availability of emotional information may also be perceived as disruptive by the learner.

To sum up, the features of an EAT can determine the frequency and effort for perceiving-monitoring the emotional information provided by other learners sharing the same computer-mediated learning environment. A useful EAT may therefore not be enough if paying attention to this information may be perceived useless or exceed learners' cognitive resources at a given time during the task.

2.5.4 From Inter-Personal Emotion to Meaning-Making

Passage #4 consists in extrapolating meaning from the emotion of other learners sharing the same computer-mediated learning environment and whose emotions have been monitored by the learner herself. This passage can be illustrated through a series of questions the learner is likely to ask herself.

The first question may be: do I understand the emotional experience of my colleagues (Eligio et al., 2012; Hall et al., 2018; Van Kleef, 2018)? This can be considered in two ways. First, in absolute and non-contextual terms. For instance, what does it mean to be *confused*, *happy*, *ashamed*, . . . for someone else? Or what does it mean to be high vs. low on the *Valence*, the *Arousal,* . . . dimensions for someone else? Is it the same as for me or not? Second, in relative and contextual terms. That is, what does it mean that my colleague is *confused*, *happy*, *ashamed*, . . . in this specific situation? Or what does it mean that my colleague is high vs. low on the *Valence*, the *Arousal*, . . . dimensions in this specific situation?

A subsequent questions may consists in wondering: how does the emotional experience of my colleagues impact their behavior? As seen in the inter-personal function of emotion in Section 2.3, causes and consequences of behavior can be inferred from emotion in others. This passage does not only require the ability to discriminate between different emotional experiences in the abstract and in the relative terms, but also to *attach* potential effects on the learning task at hand.

The design of an EAT can intervene in this passage even more preeminently compared to passage #4 from the inter-personal perspective. In that case, the learner herself can be more confident in extrapolating meaning making from her own emotions, since she is the one who has both experienced and decided to inject them into the system. It is therefore safe to assume the learner can come closer to the *intended* meaning, since she disposes of background knowledge. On the contrary, inter-personal meaning making may vary greatly based on the amount of background knowledge, depending for instance, on how acquainted the learner is with her colleagues (Hall et al., 2018; Hareli & Hess, 2010). In the extreme case of no acquaintance, as in empirical settings, the information available through the EAT may be the only *explicit* emotional information available, for as suggested by Janssen & Bodemer (2013) in Section 1.3.4, socio-affective cues can also be derived *implicitly* by the content/collaborative space of the learning task at hand (see Figure 1.1).

Finally, as stated by EASI model depicted in Section 2.3, emotion meaning-making from an inter-personal stance depends on the learner's motivation to try to decipher the emotional information at hand. The model also suggests that this may be further influenced according to how *appropriate* the learner perceived the emotions of others: the more inappropriate the emotion, the less effort may be put in extrapolating meaning-making from it.

To sum up, extrapolating meaning-making from emotion in others require the alignment of many factors. In some of them, the specific features of the EAT may have an helper function, whereas in others, the effort and motivation of the learner can not be by-passed.

2.5.5 From Intra-Personal Emotion to Inter-Personal Emotion

Passage #5 corresponds to the acknowledged and voluntarily disclosure of one's own emotional experience to others. This passage only partially overlaps with passage #1. In fact, depending on the configuration of the EAT or the learning activity, there may be situations in which emotions are expressed by learners but not shared with others, as with the use of the EMORE-L in Gaëlle Molinari et al. (2016). One possible feature of an EAT may therefore consists in making a difference between an *expressed-per-se* versus an *expressed-per-alii* emotion.

When an emotion is knowingly shared with others, all the factors linked to the social sharing of emotion come into play (A. H. Fischer & Manstead, 2016; Parkinson, 2011; Parkinson & Manstead, 2015; Rimé, 2009). These may include strategic, meta-cognitive factors. For instance, Van Kleef et al. (2011) posit that one of the primary inter-personal function of emotion is to engender social influence, that is, modify the behavior in others by harnessing the *power* of affective phenomena. If one expresses *joy* after solving a problem, this may convey a more reassuring message about the accuracy of the solution compared to a manifestation of *boredom* or *stress*.

In a collaborative setting, the emotion may also target the content space or the relational space of the learning activity (Janssen & Bodemer, 2013), as well as any *category* of emotion implicated in learning contexts: achievement, epistemic, topic or social emotions (Pekrun & Linnenbrink-Garcia, 2014b). Strategically, learner can attempt to target the disclosure of her emotional experience exclusively or preeminently toward a space or *category* of emotions, intentionally avoiding others. For instance, one can disclose frustration towards the learning material, but inhibit the same kind of manifestation towards the quality of interaction with her colleagues.

An unavoidable topic on explicit and voluntarily self-disclosure of emotion is that of emotion authenticity (De Sousa, 1987; Salmela, 2005; Scarantino & de Sousa, 2021), which often intertwines with concepts such as spontaneity, sincerity, appropriateness, and the like. From the point of view of an emotion-as-interaction approach defended in this contribution, a naive interpretation of emotion authenticity is not a necessary requirement for emotional awareness to emerge. If one *really* feels *angry*, express *anger* whereas she actually feels *frustration*, or express *anger* as a strategic signal – "*be aware, this can make me (or someone else in the group) angry*" – they all comes down to the same symbolic element to manipulate in self-regulation, co-regulation or socially shared regulation of learning. On the other hand, if emotional expression is voided of any *true* affective experience, then it becomes only a codified system of signals according to the definition of the emotional terms adopted (Scarantino, 2017; Scarantino & Griffiths, 2011).

Even though a voluntarily self-report EAT can not even scratch the surface of the complexity of emotion authenticity in the sense of correspondence *to the truth*, its defining features can make *internal authenticity* more or less manifest. For instance, if one is expected to express only a discrete emotion term, the *internal authenticity* reposes on only one unit of information. If one has to provide dimensional evaluation of some form of criteria, plus the discrete emotions term, the *internal authenticity* is augmented by the synergy of more than one piece of information. In other words, to *fake anger*, one has to go through more trouble in expressing a cohesive message that denotes what *anger* should be in a *truly angry* agent. By contrast, if we believe in learners' good-faith, the features of the EAT may precisely help them make the message more *authentic*, in the sense that it better conveys the felt experience. In this regard, an EAT that provides learners with a better experience in expressing how they feel could foster emotion disclosure even in reluctant learners.

To sum up, the disclosure of emotion, which technically means that an emotion from intra-personal becomes inter-personal on the EAT interface, can be sustained by an EAT at least by making the experience of expressing an emotion as valuable as possible. This mainly means providing learners with the possibility to *encode* in an expressed emotion information that they consider to have strategic potential.

2.5.6 From Inter-Personal Emotion to Intra-Personal Emotion

Passage #6 concerns emotions that from the inter-personal level pass to the intra-personal level. This segment can be interpreted mainly by two processes: (1) the comparison between one's own emotions and that of others (which may have been grouped also with the previous segment); and (2) a form of emotional contagion, affective reaction or meta-emotion (Barsade, 2002; Miceli & Castelfranchi, 2019; Van Kleef, 2017) that entails from knowing the emotion of others.

Emotion comparison consists in assessing the extent by which the emotion experienced by others are similar or different from the one experienced by the learner herself. As mentioned above A. H. Fischer & Manstead (2016) identify two major social functions of emotion: affiliation and distancing. Direct emotional comparison can thus sustain these fundamental functions. As it is the case with the assumption that *positive* emotions lead to learning and *negative* emotions do not, though, an equation stating that similar emotions in oneself and others lead to affiliation, whereas different emotions lead to distancing seems oversimplified. In a learning group where everybody feels constantly *irritated* by each others, learners would be more distanced than affiliated. On the contrary, a group where some learners are *stressed* and other are *confident* may lead to a reciprocal affiliation: confident learners can calm down stressed ones, whereas stressed ones can make the others less confident and more attentive to the requirements of the assignment.

Emotional comparison in itself can be performed from an analytic point of view, without triggering any affective consequence. At the same time, the emotion in others can represent triggers for the emotion in oneself, directly or indirectly through the comparison itself. As stated by theories such emotion contagion (Barsade, 2002), affective reactions (Van Kleef, 2017), or meta-emotion (Miceli & Castelfranchi, 2019) already mentioned above, emotion in others can elicit emotions in oneself. In addition, the comparison in itself can be an emotional trigger. For instance, a *proud* colleague may not in itself trigger an affective reaction, but it may do so if the learner compares with her own *disappointment*. The comparison can at this point produce *jealousy* or *bitterness* depending on, for instance, whether the *pride* is considered deserved or not.

The transition from the inter-personal to the intra-personal emotion can then flows in the intra-personal or inter-personal meaning making according to the perspective that is conferred to the information by the learner. The learner can be more interested in asking herself *what does it mean for me that the others have such and such emotion?*, taking or not into account that she has – also or instead – such and such emotion. On the other hand, the learner can be more focused on asking herself *what does it mean to others that they have such and such emotion, when I – also or instead – have such and such emotions?*. This interplay can be particularly important in collaborative settings, where this mutual-modeling process is considered an integral part of learning (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009), as mentioned in Section 1.3.4.

To sum up, the specific features of an EAT may guide learners in positioning themselves with respect to the emotion of others. How emotion is represented can accentuate the differences or the commonalities in the emotional experience in oneself and in others, as well as sustaining to different extent the reciprocal representation that learners have of each others. Some form of direct and persistent comparison can also be more prone to entail affective reactions due to the comparison itself.

2.5.7 From Meaning-Making to the Learning Activity

Finally, passage #7 coalesces all the other passages and goes back from the overarching meaning-making step to the learning activity. This is the passage that can have even more more intervening factors compared to the previous ones. As already mentioned, contrary to affect-aware systems – where some form of intelligence may intervene and guide this passage – with an awareness tool learners are bestowed with the full responsibility of fruitfully integrating the information into the task at hand. One simplified way to depict this passage is by pointing out three possible scenarios of how emotional awareness may be concretely implemented into the learning activity and, by extension, within learning processes and outcomes. Before describing the passages, though, it is important to recast that the model depicts iterative cycles in the use of an EAT: it is not the assessment of a comprehensive and definitive outcome. As the learning activity progress, there may be a number of cycles of one type, followed by cycles of another type, and so on. Each successive cycle takes into account what meaning the learner has extrapolated by combining information from emotional awareness specifically and all the other significant cues she may have retained.

The first scenario is that no meaning-making is extrapolated from emotional awareness, which entails that the course of action in the learning activity is unaffected by the use of the EAT. This may be due to a number of reasons, not all of them stemming from the inefficacy of the EAT or emotional awareness more generally. One obvious reason consists in the fact that no (new) emotion has been triggered since the last cycle. One of the reasons emotions are considered an important phenomenon is that they do not occur at all time, but rather when crucial events are at stake. Another reason for a *blank* iteration is that the available (new) information is ambiguous or unclear, so that the learner prefers to wait and not steer the course of action based on what she considers insufficient evidence. Yet another reason is that the learner has no immediate interest in wondering how the learning activity could be adapted based on the (new) emotional information available. She may be too focused on some other activity that at this particular moment may be blocking both the integration of the available information and the production or consumption of new emotional information.

The second possible scenario is that the learner decides to adapt the learning task by explicitly targeting some form of emotion regulation (J. J. Gross, 2015; Hoffmann et al., 2020; Reeck et al., 2016; Zaki & Williams, 2013). A modulation of the course of action is produced to attain specifically the emotional experience and not, for instance, the topic at hand or the negotiation of learning goals. As mentioned in the current and previous chapter, emotion regulation can apply at the individual or social levels (*ibid.*). For instance, after expressing an emotion and extrapolating the information that this very emotion is assessed as inadequate to the learner's goal, she may use one of the emotion regulation strategies (*i.e.*, situation selection, situation modification, attentional deployment, cognitive change or response modulation) to act upon her own emotional experience (J. J. Gross, 2015). If the relevant emotional information comes from another colleague sharing the same computer-mediated environment, the learner can decide to use it as a reference for her own behavior (Hareli et al., 2013; Van Kleef, 2018) or to apply emotion regulation at the inter-personal level (Reeck et al., 2016).

The last possible scenario presented here consists in a wider form of regulation of learning, identified in the triad self-regulation, co-regulation and socially-shared regulation of learning (Järvelä et al., 2015, 2016; Miller & Hadwin, 2015; Winne, 2015), illustrated in Section 1.1.3. In a sense, emotion regulation in learning settings can be viewed as a particular form of these wider strategies, which extend beyond the emotional experience. The (new) emotion information obtained through the EAT and processed through the overarching meaning-making process can be integrated into the learning task at various levels. A learner can for instance use the emotion of others as a feedback on its own contribution to a common task, duplicating the effort if those emotions denote appreciation, or assuming a more passive stance if those emotions question the quality of the input provided. The fact that the model is cyclical does not yet prevent learners to try to establish patterns in their regulatory processes. For instance, a learner could realize that she has a tendency to attribute to others emotional reactions that are diminishing of her own work (Scherer, 2021b), when in fact this is not supported by the information available through the EAT. This could lead the learner to keep this *bias* in mind and eventually work on it over time.

To sum up, this last passage represents the core of emotional awareness as instrumental to learning processes and outcomes. It defines whether endowing computer-mediated learning environments with emotional awareness can create a virtuous circle consisting in the emergence and processing of meaningful emotional information, which is then translated in concrete elements related to the learning activity. As proposed in the description of this last passage, this translation can be viewed as the alternation of of three possible scenarios: the (new) emotional information has no impact, it impacts emotion regulation, or it impacts learning regulation more broadly. Throughout the model, emotional information has been preceded by the optional (new) prefix. This was meant to consider that, in case of persistence of emotional information over time, learners can also reassess past experience under new perspectives. For example, when looking at how their emotions have evolved during a semester, learners could assess that all the *confusion* and *frustration* they have been living was worth the trouble, considering their sentiment of competence acquired at the end of the course. The *confusion* and *frustration* are not new emotional information in itself, but it becomes *new* in light of how events have turned out.

2.6 Use of the Abstract Model to Define Research's Aims

As stated in the introduction of this chapter, one of the aim of the model is precisely to evoke a sort of *domino effect*: all the techno-pedagogical pieces must align for emotional awareness to be effective. At the same time, it is possible to break down the whole mechanism in what may be more manageable parts to deal with, especially from the point of view of research in the domain.

The model can thus be adopted as a guideline to think about and communicate what part of emotional awareness is (primarily) aimed by a contribution. This does not mean that the whole process is beyond reach, but it presupposes that for studying the whole process, more *pieces* must be accounted for – theoretically, technically and/or empirically. For instance, the thesis aims at investigate consistently passages 1 and 3, with 5 and 6 also implicated from a mutual-modeling perspective. On the other hand, passages 2 and 4 are considered with a more exploratory outlook. Finally, passage 7 is considered beyond the thesis's reach altogether.

By breaking down the all process, it is also possible to pinpoint what causal mechanism can be responsible for the outcome of that specific part(s). For instance, self-report can be instrumental in passages 1 and 3, by forcing learners to reflect on their own emotional experience. It also play a prominent role in passage 5, for it underpins the disclosure of the emotional information to others. The implementation of an emotion structure is a good example of an overarching causal mechanism, which is meant to mediate or moderate the whole gamut of sub-processes implicated in emotional awareness. In this regard, the next chapter of the thesis will define precisely what is meant by an emotion structure, and how this structure can intervene in producing and consuming emotional awareness. Furthermore, the empirical contributions of Part III will refer to the abstract model to explain their specific aim, so more concrete examples will be provided.

2.7 Summary

This chapter introduced the concept of emotional awareness, around which emotion awareness tools are built. From the broad definition of the concept, three underlying assumptions have been derived. The first two assumptions concern the role that emotion play at the intra-personal and the inter-personal levels respectively. The third assumption posits that emotional awareness can be fruitfully conveyed in a computer-mediated learning environment, by harnessing for instance the power of computational devices. For each assumption, an overview of theoretical and empirical work has been provided, in an attempt to sketch the complexity and variety of approaches on the subject. In order to break down this heterogeneity in a reasoned organisation, an abstract model of emotional awareness, consisting in seven integrated passages, has been outlined. The abstract model can be used as an instrument to think and communicate about research on the topic. In this regard, the specific aims of the thesis have been used in a preliminary way to illustrate the model.

Chapter 3

Defining the *Unit* of Measure to Self-Report and Convey Emotional Awareness

This chapter lays out the foundation for implementing the emotion structure into an EAT. The structure of an emotion is determined by its definition. For this reason, the chapter starts by highlighting the difficulties in reaching a consensual definition of what an emotion is. Next, a brief outline of competing emotion theories is provided. Among these theories, the contribution adopts appraisal theories of emotion, whose characterizing features are depicted in the third section of the chapter. Among appraisal theories, the Component Process Model, depicted in section 4, is retained as the theory of reference for the implementation of the EAT. The framework presents also the advantages of proposing a concrete link with emotional competence, which is discussed in the last section of the chapter.

3.1 What is an Emotion?

What is an emotion? is a question that literally appears in a multitude of scientific contributions. For instance, it is the title of a seminal article by William James (1884), which, despite being more than a century old, is still source of lively debate (Ellsworth, 1994; Scherer, 2005). There are many iconic statements that testify to the difficulty of studying emotions from a scientific perspective. Fehr and Russel (1984, p. 464) argued that "[e]veryone knows what an emotion is, until asked to give a definition", which resonates with Kleinginna and Kleinginna (1981), who, around the same time forty years ago, collected almost one hundred definitions. More recently, Frijda and Scherer (2009, p. 142) still confirmed that "emotion may be one of the fuzziest concepts in all of the sciences", which was empirically corroborated by Izard (2010), who attempted – and failed – to extrapolate a unitary definition from the responses of 35 leading researchers from different fields to 6 questions about emotion definition, function, activation and regulation.

In the last few decades, though, a growing consensus has emerged over a *minimal*, consensual definition of some characteristics of emotion that are shared by different scholars and models of emotion (Frijda & Scherer, 2009; Kleinginna & Kleinginna, 1981; Moors, 2009; Sander, 2013). Sander (2013, p. 16) resumes this consensus in four points: (1) emotions are multicomponent phenomena, spanning affective, cognitive, behavioral and physiological processes ; (2) emotions are a two-step process involving *emotion elicitation* mechanisms that produce *emotional responses*; (3) emotions have relevant object focus, being triggered by internal or external events/stimuli that are relevant for the organism; and (4) emotions have a brief duration compared with other affective phenomena such as moods, attitudes or preferences.

Within this consensual framework, though, different positions persist, which have both theoretical and applied consequences in the study of emotion *per se*, as well as in conjunction with other phenomena. These differences are often rooted in historical, cultural and epistemological contexts, within and across disciplines and approaches – see Dixon (2003) and Plamper (2015) for overarching perspectives, or Sander (2013) for a more concise introduction. Some of the broader controversies regarding emotion can be encapsulated in onion layers that are highly interconnected and mutually influenced. From the outset to the inset, controversies may concern:

- How emotion relates to, and differ from other phenomenon of the mind/body such as intelligence, rationality, consciousness or various forms of cognition (Bechara, 2004; Brosch et al., 2013; Leventhal & Scherer, 1987; Lieberman, 2019b; Murphy & Hall, 2011; Pessoa, 2013; Rolls, 2014; Salovey & Mayer, 1990);
- How emotion relates to, and differ from other affective phenomena such as moods, preferences, attitudes, motivation, desires, feelings, sentiments, or passions (Colombetti, 2014; Dixon, 2003; Linnenbrink-Garcia et al., 2016; Scherer, 2005; Schwarz & Clore, 2003);
- Whether there are, and what are the building blocks and features covering the whole process, from causation to consequences, shared by all phenomena identified as emotion (Adolphs & Anderson, 2018; J. J. Gross, 2002; Moors, 2009; James A. Russell, 2003; C. Smith & Ellsworth, 1985);
- Whether there are a countable number of, and how to distinguish between discrete emotions – such as *anger*, *fear*, or *happiness* – and if is it possible to categorize them in taxonomies (Ekman, 1992; Fehr & Russell, 1984; Fontaine & Scherer, 2013; Pekrun & Linnenbrink-Garcia, 2014a; James A. Russell, 2003; Sander, 2013; Scarantino & Griffiths, 2011)
- How a specific emotion, assuming that a discrete emotion exists, should be referred to – both from an everyday communication and scientific conceptualization points of view – considering differences between cultures and languages (Douglas-Cowie, Campbell, Cowie, & Roach, 2003; Ekman & Cordaro, 2011; Fontaine et al., 2013; Gillioz et al., 2016; Grandjean et al., 2008; Scherer, 2013a; Torre & Lieberman, 2018).

The implementation of an EAT, especially one that has a multipurpose vocation, must therefore face the complexity of defining the *unit* that it aims to make people aware of (Boehner et al., 2007). In this regard, a first, broad decision that shall be considered can be related to what Scarantino and de Sousa (2021) identify as the two *desiderata* of emotion definition. The first consists in shaping the definition of emotion in order to achieve compatibility with ordinary linguistic usage, which corresponds to a descriptive definition. The second aims at finding a definition of emotion that is theoretically fruitful, that is, a prescriptive definition. The two desiderata are not mutually exclusive, since even a prescriptive definition should try to maintain as much compatibility with ordinary linguistic usage as possible (Scherer, 2005). But, in the case of incompatibility with ordinary intuition, a prescriptive definition should favor theoretical generalization (Scarantino, 2012). Roughly translated in more concrete terms, a definition of emotion should guide the implementation of an EAT in deciding whether it shall convey what users consider an emotion to be from a *folk* perspective, or rather what an emotion should be considered according to a *scientific* perspective (Fontaine & Scherer, 2013; Ogarkova, 2013; Scarantino, 2012; Scherer, 2005).

From a pure usability standpoint, that is, the perspective of a learner that must face a tool she is not used to have in a learning environment, a *folk* perspective should not be discarded as simplistic. For instance, an EAT could be based on the most frequent affective states encountered by learners (D'Mello, 2013; Gaëlle Molinari, Chanel, et al., 2013; Reis et al., 2018), without any relationship with a prescriptive definition, and regardless of whether these states are more or less prototypical of emotion (Fehr & Russell, 1984; Scarantino & Griffiths, 2011). Even more bluntly, an EAT could avoid the problem altogether by prompting users to share *what they think an emotion is*, letting them construe the meaning of how they feel without any guidance (Barrett, 2006; Barrett, Gross, Christensen, & Benvenuto, 2001; James A. Russell, 2003). On the other hand, from a multipurpose point of view – which is in and on itself a form of generalization – the link with a *scientific*, theory-driven perspective presents several advantages.

First, by mapping the tool to a prescriptive definition, emotion not only emerges according to what it *is*, but also according to what it *does* and why. As seen in the previous chapter, there is a growing body of literature focusing on the functional role emotion plays both at the intra- and inter-personal levels (Adolphs & Andler, 2018; A. H. Fischer & Manstead, 2016; Keltner & Haidt, 1999; Levenson, 1999; Leventhal & Scherer, 1987). Thus, a theory-driven perspective may contribute to better assemble these functions and take full advantage of them in the implementation of the EAT (Brosch & Sander, 2013; Pekrun, 2005; Van Kleef, 2018).

Second, tailoring an EAT on emotion theory may contribute to guide learners in discovering and using it, since they can be both interested and reassured by identifying underlying mechanisms that are thought to be an integral part of the phenomenon (Veronique Tran, Páez, & Sánchez, 2012). The fact that a dedicated tool is provided in the environment for some *corollary* activities does not automatically mean that it will be adopted (Kreijns et al., 2003).

Third, from a research perspective, linking the EAT to a theory facilitates the assessment and comparison of measures collected through or about it, and therefore foster common ground in the contributions that shall decide to adopt it (Izard, 2010; Scherer, 2005). In the meantime, the EAT would also benefit from the upcoming theoretical and applied contributions in the literature that relate to the shared emotion theory, ensuring the tool can progress over time, or otherwise be dismissed if some competing theory provides a better alternative (Ferguson & Heene, 2012; Meehl, 1990).

These reasons advocate the adoption of a theory-driven approach to the definition of the emotion *unit* to be implemented into the EAT. The next section will thus address the topic of emotion theories.

3.2 Theories of Emotion

As one may expect from the complexity in defining an emotion outlined in the previous section, there are a multitude of emotion theories and taxonomies that try to categorize them (see for example Moors, 2009; Sander, 2013; Scarantino & de Sousa, 2021). Adopting a perspective spanning philosophy and affective sciences, Scarantino and de Sousa (2021) identify three traditions in which most of emotion theories may belong: (1) the Feeling Tradition, which considers the conscious experience of an emotion as the fundamental characteristic; (2) the Evaluative Tradition, which focuses on how emotion consists in, or is crucially defined by the way eliciting events are evaluated; and (3) the Motivational Tradition, which identifies emotion with specific internal states driven by want or need to satisfy a goal.

In psychology more specifically, the last decades have been characterized mainly by three major families of emotion theories that account for the *whole* emotional phenomenon (*i.e.*, elicitation and response): (1) basic emotion theories (Ekman, 1992; Ekman & Cordaro, 2011; Hutto, Robertson, & Kirchhoff, 2018; Keltner, 2019), stemming from the assumption that there exist a fixed number of discrete, *primary* emotions, each of them presenting a fixed pattern of elicitation and response (2) constructivist/dimensional emotion theories (James A. Russell, 2003, 2009; Stanley & Meyer, 2009), driven from a dimensional approach to emotion, in which a persistent core affect is determined by a bi-polar combination of physiological arousal (*e.g.*, activating vs. deactivating) and psychological assessment of the perceived valence of events (*e.g.*. pleasant vs. unpleasant); and (3) appraisal emotion theories (Moors et al., 2013; Scherer, 2009b; C. Smith & Ellsworth, 1985), in which discrete emotion are determined by the ongoing, cognitive evaluation of events on a number of (dimensional) criteria. These three families emerged from different historical backgrounds, and their foundation are often rooted in philosophical perspectives with secular traditions. In *psychological* era, we can trace back seminal works on the three families in the second half of the last century (Arnold, 1960; Lazarus, 1966; James A. Russell, 1980; Scherer, 1982). Since then, theories have influenced each other, converging on some points – most notably the fact that emotion is not unitary but comprises different components – and maintaining different positions on other points, such as eliciting and differentiation mechanisms in discrete emotions (Frijda & Scherer, 2009; Sander, 2013). New and often interdisciplinary theories that have emerged lately resume and extend some ideas of one or more of these families. It is the case, for instance, of the *enacting* approach proposed by Colombetti (2014), which bestow to emotion, from a phenomenological perspective, a more *active* role compared to the *passive* response to events that is sometimes implied by some emotion theories. Another example is represented by the *theory of affective pragmatics* proposed by Scarantino (2017), which pursue the idea of the existence of emotion as natural kinds, as in basic emotion theories. In other cases, such as in the *theory of constructed emotion* advocated by Feldman Barrett (Barrett, 2006, 2017, 2018), there is considerable discontinuity with previous work, challenging the principle according to which emotions are responses or exist as natural kinds. For instance, Feldman Barret (2018, pp. vii–viii) argues that "[e]motion are real, but not in the objective sense that molecules or neurons are real. They are real in the same sense that money is real – that is, hardly an illusion, but a product of human agreement". Some contributions have also attempted to highlight common grounds rather than distinction between theories (Scherer, 2022), or adopt a theory-independent perspective and rather focus on empirical contributions by investigating the accuracy in making predictions about the emotion experienced by a person in a controlled environment (Scherer & Moors, 2019).

A second important choice in the implementation of a theory-driven EAT is therefore to define which emotion theory or set of theories it relates to. At present, no emotion theory provides clear and undisputed supremacy in terms of explanatory power of what an emotion is (Moors, 2009; Sander, 2013; Scarantino, 2018; Scarantino & de Sousa, 2021; Scherer, 2022). As a consequence, the choice of the theory may be assessed rather in functional terms: which tradition/theory may provide the best framework to foster emotional awareness? The present contribution makes the opinionated assumption that the Evaluative Tradition, and more specifically appraisal theories of emotion, may best fit this purpose. It is posited by this choice that an EAT may convey instrumental information to learners, if that information is crucially determined by how learners construe emotion meaning from the situation at hand, an assumption also shared by other scholars investigating the relationship between emotion and learning (Lavoué et al., 2020; Gaëlle Molinari, Chanel, et al., 2013; Pekrun, 2006; Shuman & Scherer, 2014). The next section offers an overview of this family of emotion theories.

3.3 Appraisal Theories of Emotion

Appraisal theories of emotion originated in the 1960's with the work of Arnold (1960) and Lazarus (1966) and were then formalized in the 1980's by independent scholars reaching similar conclusions about the need to explain how different emotions may be elicited by the same event, for different persons and in different situations (Roseman & Smith, 2001; Scherer, 1982; Schorr, 2001; Siemer et al., 2007; C. Smith & Ellsworth, 1985). The core element shared by theories belonging to the appraisal family is resumed by the premise that emotions "are adaptive responses which reflect appraisals of features of the environment that are significant for the organism's well-being" (Moors et al., 2013, p. 119), where well-being roughly encompasses the satisfaction or obstruction of someone's wants, needs or goals that are considered of major concern (Frijda, 1986; Roseman & Smith, 2001; Scherer, 2005). This feature is nevertheless not unique to appraisal theories, since other theories reckon adaptive functions to emotion or motivational implications (Adolphs & Anderson, 2018; Ekman, 1992; Scarantino, 2017; Scarantino & de Sousa, 2021). Appraisal theories are therefore more thoroughly characterized by the dynamic, recurrent and ongoing nature they confer to emotion in sustaining the adaptive response (Brosch & Sander, 2013; Oatley & Johnson-Laird, 2014; Roseman & Smith, 2001; Scherer, 2009b; Schorr, 2001; Siemer et al., 2007). It is in fact common among appraisal theories to consider emotion as a process, rather than a state (*ibid.*). The bounded process unfolds over brief periods of time and can be modified in light of new information coming from any of the organismic subsystems or components, which include "an appraisal component with evaluations of the environment and the person-environment interaction; a motivational component with action tendencies or other forms of action readiness; a somatic component with peripheral physiological responses; a motor component with expressive and instrumental behavior; and a feeling component with subjective experience or feelings" (Moors et al., 2013, pp. 119–120).

An example of an appraisal theory is Pekrun's Control-Value theory of achievement emotions (Pekrun, 2006), already sketched in Section 1.1.1. In Pekrun's own words (*ibid*, p. 317):

> The control-value theory described here posits that two groups of appraisals are of specific relevance for achievement emotions: (1) *subjective control* over achievement activities and their outcomes (e.g., expectations that persistence at studying can be enacted, and that it will lead to success); and (2) the *subjective values* of these activities and outcomes (e.g., the perceived importance of success).

What emotion a learner will feel in an achievement context, such as studying for a test, will depend on the specific pattern emerging from the combination of appraising the subjective control and the subjective value. For instance, "[i]f an achievement activity (e.g., studying) and the material to which it relates (e.g., learning material) are positively valued, and if the activity is perceived as being sufficiently controllable by the self, *enjoyment* is assumed to be instigated" (*ibid.*, p. 323). Conversely, "[i]f the activity is perceived as being controllable, but is negatively valued (e.g., when effort required by the activity is experienced as aversive), *anger* is expected to be aroused. If the activity is not sufficiently controllable, *frustration* will be experienced" (*ibid*, p. 323, all italics in the text).

According to the dynamic, recursive and ongoing nature of emotion posited by appraisal theories, then, the evaluation of the activity may change instantly. For instance, the learning material may be perceived as too difficult at first, lowering learners' subjective control, but become clearer after a synthesis or a figure. Likewise, the subjective value may suddenly arise if, after reading a few pages useless to one's needs, the learner finds a passage she can directly apply to solve a problem that has been bothering her ever since. Conversely, the very same passage may be considered of scarce value from her colleague, who has already solved the problem beforehand.

The Control-Value theory of achievement emotions may therefore seem to be a valid candidate to guide the implementation of an appraisal-driven EAT, but it only partially comply with the multipurpose vocation. In fact, learning is characterized by different *kinds* of emotion, which comprise, but are not limited to, achievement emotions. As illustrated in Section 1.1.1, Pekrun and Linnerbrink-Garcia (2014a) reckon that learning also triggers epistemic emotions, topic emotions, and social emotions. An overarching appraisal theory would therefore be more suitable to a multipurpose EAT. In this regard, appraisal theories, even if they share considerable overlapping, differ on several key points, which are enumerated in a state of the art proposed by Moors et al. (2013). Most of the points are too detailed for the present purpose and the complete list would lack instrumental information to be fully explained. Only the points concerning appraisal criteria are considered to be of primary relevance and will thus be extrapolated from the aforementioned source.

As a reminder, appraisal is considered a process that scans events outside or inside the organism and evaluates them against a set of criteria related to the organism's major concern. One of the main source of divergence in appraisal theories refers, precisely, to the number of appraisal criteria the event is evaluated against, the order of evaluation, as well as the possible outcomes of every criterion. An appraisal criterion can have dimensional, ordinal or categorical values. For instance, the Control-Value theory described above (Pekrun, 2006) considers that the two appraisals Control and Value can assume ordinal values *none*, *low*, *medium* and *high*. Other appraisal theories postulate dimensional criteria, with potentially infinite values along a continuum, or categorical criteria, with a predefined set of possible values. As stated by Moors et al. (2013):

> The number and nature of the appraisal variables and/or values is closely related to the number and nature of the emotions that one can or wishes to explain. In general, more emotions require more appraisal variables and/or more appraisal values. Turning it around, more appraisal variables and/or more appraisal values allow more variety in emotions. Two appraisal variables with two values each can account for four emotions. Seven appraisal variables with an infinite number of values each can account for an infinite number of emotions. The number and nature of the emotions that one wishes to explain can be traced back to metatheoretical choices such as whether one strives for parsimony and/or a focus on natural language descriptors of emotions, on the one hand, or exhaustiveness and/or a focus on variety, on the other hand.
> — Moors et al. (2013), p. 121-122

This contribution adopts the Component Process Model theoretical framework (Scherer, 2001, 2005, 2009b, 2013b) as the emotion theory of reference for reasons that will be illustrated in the next section, alongside a description of the model.

3.4 The Component Process Model Theoretical Framework

The Component Process Model (CPM) is a theoretical framework bridging appraisal and componential approaches to emotion, first proposed by Scherer (1982) in the 1980's and then refined over the years (2001, 2005, 2009b, 2013b). It has been applied to different contexts: the reader may find discussions of the model in relation with learning and education in Shuman & Scherer (2014), with affective computing in Scherer et al. (2010), with emotional meaning-making in Fontaine et al. (2013), with emotional competence/intelligence in Scherer (2007), and with neuroscience in Sander, Grandjean, & Scherer (2018). An overarching and accessible description of the model can be found in Scherer (2005), whereas Scherer (2001) and Scherer (2009b) are more focused on the dynamic, recurrent and ongoing features of the model. Finally, more recent developments are discussed in Scherer (2013b), Scherer (2019a) and Scherer (2022). The choice of the CPM as the main theoretical source for the implementation of an emotion structure into an EAT is warranted by the following reasons.

First, the model encompasses the whole process of emotion elicitation, manifestation, and representation in oneself (Sander, 2013) and can be extended to mechanisms in others (A. H. Fischer & Manstead, 2016; Rimé, 2009; Van Kleef, 2018). It is therefore consistent both with the intra-personal assumption of emotional awareness illustrated in Section 2.2, and with the inter-personal assumption depicted in Section 2.2.

Second, the model draws a clear distinction between different affective phenomena. This is particularly relevant for contextual reasons. Pekrun and Linnenbrink-Garcia (Pekrun & Linnenbrink-Garcia, 2014b, p. 2), namely, point out that "[i]n the broader education literature, affect is often used to denote a broad variety of noncognitive constructs including emotion but also including self-concept, beliefs, motivation, and so on [. . .]. In contrast, in emotion research, affect refers to emotions and moods more specifically". As depicted in Figure 3.1, the CPM considers emotion: strongly event focused and driven by appraisal (i.e., elicited by an event that can be identified and evaluated), of relatively brief duration, subject to abrupt changes, and with the potential of being regulated.

Types of affective states/dispositions: (examples in parentheses)	**Emotions** (e.g. angry, sad, joyful, fearful, ashamed, proud, elated, awestruck, desperate)	**Moods** (e.g. cheerful, gloomy, irritable, listless, depressed, buoyant)	**Affective interpersonal stances** (e.g. distant, cold, warm, supportive, contemptuous)	**Attitudes** (e.g. liking, loving, hating, valueing, desiring)	**Affective personality dispositions** (e.g. nervous, anxious, reckless, morose, hostile, envious, jealous)
Event focus	+++	+ → ++	+	0	0
Appraisal elicitation	+++	+	+	+	0
Action tendencies	+ → +++	+	++	+	+
Physiological responses	+ → +++	+	0 → ++	0 → +	0 → +
Motor Expression	+ → +++	+	0 → ++	0 → ++	0 → +
Component synchronisation	++ → +++	+	+	0	0
Feeling	+++	+ → ++	+ → ++	+ → ++	0
Verbalization	0 → ++	0 → +++	0 → ++	0 → +	0 → +
Intensity	++ → +++	+ → ++	+ → ++	0 → ++	0 → +
Duration	+	++	+ → +++	++ → +++	+++
Rapidity of change	+++	+ → ++	+	0 → +	0
Regulation potential	+ → ++	+ → ++	+	0 → ++	0

Note: Occurrence/Strength: xxx often/strong, xx sometimes/medium, x rarely/weak, 0 never/absent; → range

Figure 3.1: (ref:emtion-vs-other-affective-phenomena-caption)

Third, the model bestow to the subjective experience of the emotional episode a central role in extrapolating meaning-making both at the individual and interpersonal levels (Grandjean et al., 2008; Scherer, 2005). This is consistent with an interactional approach to emotion (Boehner et al., 2007) illustrated in Section 1.2.2.

Forth, the CPM can be related to emotional intelligence/competence in concrete, practical terms (Scherer, 2007). As illustrated in Section 1.1.3, socio-affective competences are one of the central areas of research bridging affective sciences and learning/education sciences.

Finally, the model is computational in nature (Scherer et al., 2010). A computational model of emotion (Marsella, Gratch, & Petta, 2010b) may be helpful in complying with the third main assumption of emotional awareness of Section 2.4, namely that emotional awareness can be fruitfully conveyed in a computer-mediated learning environment.

This section first proposes an overview of the model and its components according to the objectives of the thesis by combining a computational (Scherer, 2010b), emotion meaning-making (Scherer, 2013a; Scherer & Fontaine, 2018) and voluntary self-report (Scherer, 2005; Scherer, Shuman, et al., 2013) perspectives. Then, it provides further information on 3 *modules* that can be interpreted in an Input/Output process.

3.4.1 Overall Description of the Model

The CPM adopts a functional perspective according to which "[e]motions have developed in the course of evolution to replace rigid instincts or stimulus-response chains by a mechanism that allows flexible adaptation to environmental contingencies by decoupling stimulus and response, creating a latency time for response optimization" (Scherer, 2010b, p. 48). Response optimization emerges as the result from the synchronization of five organismic subsystems performing specific functions. The state of each subsystem during an emotion episode corresponds to one of the five components of the model, namely: (1) the cognitive, (2) the neurophysiological, (3) the motivational, (4) the motor expression, and (5) the subjective feeling components. According to the CPM, thus, an emotion is defined as "*an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism*" (Scherer, 2005, p. 697, italic in the text). The emotional experience therefore starts when at least three of the subsystems reach a certain threshold of synchronization, and ends when the synchronization falls below this level again.

The CPM belongs to the appraisal family of emotion theories because it states that the coordination and synchronization of the components during an emotion episode is driven by the cognitive evaluation of the situation, that is, by how the situation is subjectively appraised by the person (Scherer, 2005). The five components intertwine so that a variation in one fo them may lead to modifications in others, modifying thus the net effect of the coordination and synchronization. As stated by Scherer (2013a):

> Briefly put, the CPM suggests that the event and its consequences are appraised with a set of criteria on multiple levels of processing, producing a motivational effect or action tendency that often changes or at least modifies the status quo. Specifically, the appraisal results and the concomitant motivational changes will produce efferent effects in the autonomic nervous system (in the form of somatovisceral changes), and motor expressions are centrally represented and constantly fused in a multimodal integration area (with continuous updating as events and appraisals change). Parts of this central integrated representation may then become conscious and subject to assignment to fuzzy emotion categories, as well as being labeled with emotion words, expressions, or metaphors
> — Scherer (2013a), p. 13

The process can be understood in terms of input/output when the model is organized in three modules, as graphically represented in Figure 3.2, adapted from Scherer (2010b):

1. the appraisal module

2. the response-patterning module
3. the integration/categorization module

From the left-hand side of the model, an event, behavior or situation (the input) is subject to a multilevel appraisal; it is cognitively evaluated against a number of appraisal criteria. This evaluation leads to the patterning of a response, which comprises three out of the five organismic subsystems: (a) the motivational component, which prepares the organism to action, even if this action is not necessarily executed; (b) the neurophysiological component, which orchestrates the allocation of resources according to the action tendency; and (c) the motor expression component, which is responsible of the emergence of the underlying process to the *surface* via efferent manifestations, providing communicative cues, such as facial expressions, vocal prosody, or body postures. Finally, in the integration/categorization module on the right-hand side, the whole process – that is, the net effect of the appraisal and response-patterning – is integrated in a central representation, the feeling, of which the person may or may not be conscious. If she is conscious, the person may recur to her knowledge about features of emotion or her emotional lexicon in an effort to extrapolate a symbolic representation of the feeling, which in that case becomes the subjective feeling: the conscious output of the whole process.

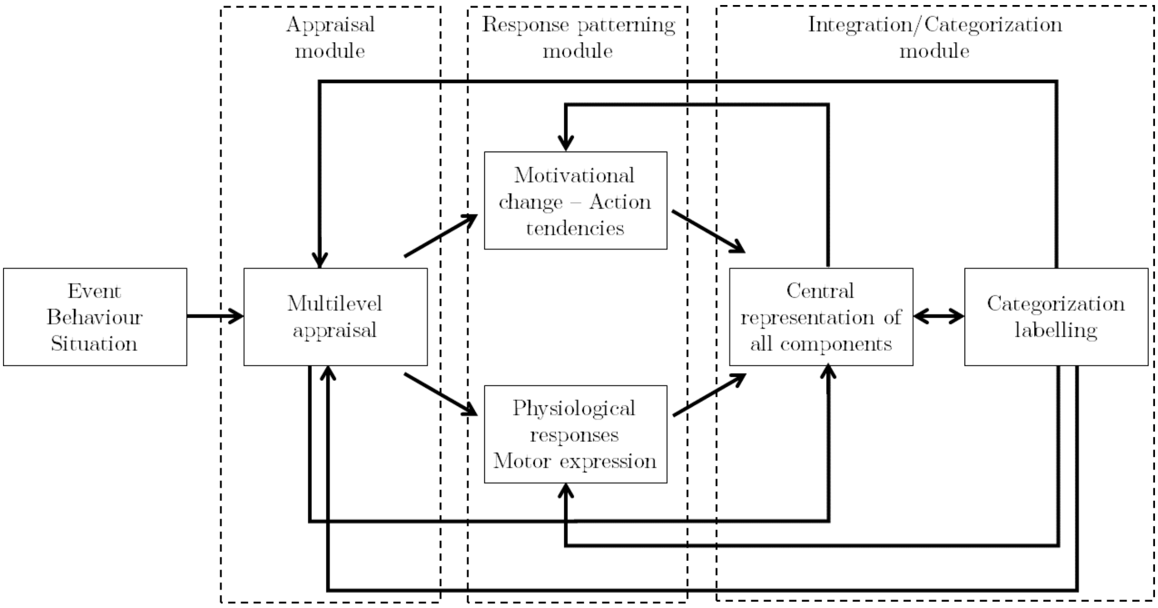


Figure 3.2: The dynamic architecture of the component process model. Adapted, with minor labeling and graphical modifications, from the original figure 2.1.1 in Scherer (2010b), p. 50.

After a brief overview of the overall model, the next subsections focus on the three modules respectively.

3.4.2 The Appraisal Module: Appraisals as Sequential Evaluation Checks

The appraisal module may be considered the most characteristic of the CPM for two intertwined reasons. First, it bridges cognition and emotion in a functional perspective (Leventhal & Scherer, 1987), avoiding thus the pitfalls of the opposition between affect and cognition, which recent developments in affective science are progressively dismissing in favor of interacting or even integrating perspectives (Dukes et al., 2021; Pessoa, 2013). Second, it plays a pivotal role in differentiating the CPM from other appraisal theories, for the CPM makes bold assumptions about how distinctive appraisal patterns lead to distinctive emotional episodes (Scherer, 1993, 2009b; Scherer & Meuleman, 2013; Scherer & Moors, 2019). In other words, the model explicitly assumes a causal mechanism instantiated by the appraisal module, which determines the unfolding emotional episode in terms of response-patterning and integration/categorization.

The CPM's appraisal module is built around the central notion of Sequential Evaluation Checks (SECs): a set of appraisal criteria the event, situation or behavior is evaluated against (Scherer, 2001, 2009b, 2013b). According to the CPM, these appraisal criteria may not only be enumerated, but also determined with respect to their timing in achieving *preliminary closure* in information processing, that is, before the result of the evaluation can be integrated by other SECs or components of the model in the unfolding emotional process (Scherer, 2013b). The CPM identifies a total of 13 SECs, listed in Table 3.1, which can be grouped in correspondence with four major appraisal objectives in adapting to a given event (Scherer, 2009b, 2013b):

- *Relevance*: how relevant is this event for me?
- *Implications/Consequences*: what are the implications or consequences of this event and how do they affect my well-being and my immediate or long-term goals?
- *Coping potential*: how well can I cope with or adjust to these consequences?
- *Normative significance*: what is the significance of this event for my self-concept and for social norms and value?

Table 3.1: Stimulus evaluation checks, organized in four groups, illustrated with typical features describing the event or the effects on the person. Retrieved from original Table 1 in Scherer (2013b), p. 151

Stimulus evaluation checks	Event or behavior/person
Relevance	
Novelty	Event is sudden, familiar, unpredictable
Intrinsic pleasantness	Event is in itself un/pleasant for the person
Goal/need pertinence	Event is important and relevant for person's goals or needs
Implications/consequences	
Causal attribution	Event was caused by the person's own/somebody else's behavior/chance; caused un/intentionally
Outcome probability	Consequences of the event are predictable
Discrepancy from expectation	Event confirmed/is inconsistent with expectations
Goal/need conduciveness	Consequences of the event are positive/negative for person
Urgency	Event required an immediate response
Coping potential	
Control	Person can control the consequences of the event
Power	Person has power over the consequences of the event
Adjustment	Person can live with the consequences of the event
Norm compatibility	
Internal standards	Event incongruent with own standards and self-ideals
External standards	Event violated laws or socially accepted norms

The evaluation of SECs can be executed at four levels of processing, which can be ranked with respect to the complexity, consciousness of the undergoing evaluation, and involvement of other phenomena or processes (Scherer, 2013b): (a) a sensorimotor level, (b) a schematic level, (c) an association level, and (d) a conceptual level. Of the four levels, the last two better comply with the use of an EAT because they are more related with higher-order cognitive processing. In particular, the conceptual level requires consciousness of the underlying evaluation and effort-full calculation in the pre-frontal cortex (*ibid.*), which are aligned with the view of an EAT as a self-reflecting and emotional experience enhancer.

Furthermore, appraisal criteria are also bidirectionally linked with cognitive functions such as attention, memory, and reasoning, as well as motivational factors and self-concepts (Sander, Grandjean, & Scherer, 2005), as illustrated in Figure 3.3. As a consequence, the specific appraisal pattern that emerges is determined by situational factors, but also by more stable structures that can lead to certain patterns to emerge more or less frequently, for instance the tendency of a student to avoid confrontation of point of views as a result of past experiences (Scherer, 2021a). In the meantime, the fact that many elements concur in determining the specific appraisal profile implies that an evaluation may cause the emotional episode to derail from adaptive binaries and lead to inefficient or even harming responses to events, as in the case of excessive anxiety or fear in taking an exam (Pekrun, 2006; Shuman & Scherer, 2014). The CPM also implies that the emotion components have links with cognitive processes (*i.e.*, the red *backward* arrows in Figure 3.3, which contributes to explain why emotion may influence for instance perception, attention, memory and decision-making (*e.g.*, Brosch et al., 2013) – that is, processes also pivotal to learning.

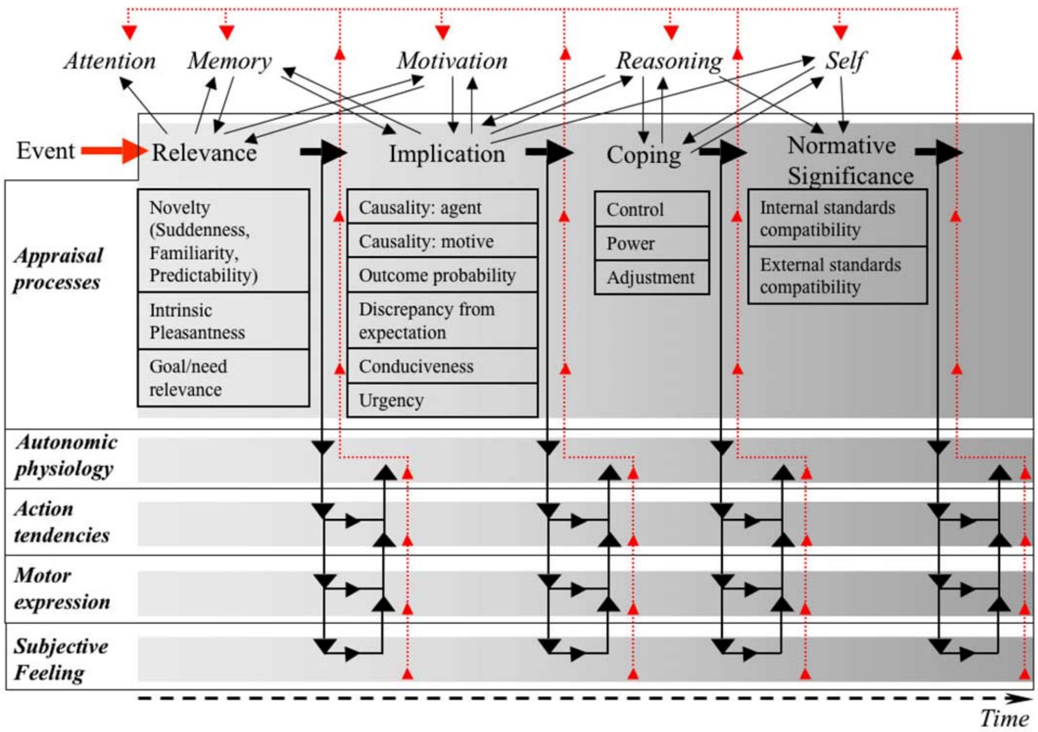


Figure 3.3: Comprehensive illustration of the component process model of emotion. Retrieved from the original Figure 1 in Sander et al. (2005), p. 321. Please note that the *conducivness* SEC does not appear in the more recent list provided by Scherer (2013b) and illustrated above.)

To sum up, in the CPM appraisal criteria play a prominent role in emotion elicitation and differentiation. As a consequence, an EAT could leverage on the pivotal role of SECs to foster emotional awareness, for instance advocating the cognitive evaluation of the salient event. In other words, the EAT could contribute to make sure that emotion is assessed at the conceptual level – the one implicating consciousness and effort-full calculations – as a means to extrapolate meaning-making from the adaptive objectives of appraisal criteria.

3.4.3 The Response-Patterning Module

The response-patterning module build upon information provided by sequential evaluation checks (SECs) described in previous subsection and is responsible for the e*mbodiement* of emotion in terms of adaptation (Scherer & Moors, 2019). Adaptive reactions can be designated at two levels: (1) intra-individual, with the regulation of the organism's functions (neurophysiological component) and the action tendencies (motivational component); and

(2) inter-individual, through the motor expression component, which comprises efferent manifestations such as facial expression, vocal prosody or body posture playing an important communicative role (see Hyniewska, Sato, Kaiser, & Pelachaud, 2019; Scherer & Grandjean, 2007a; Shuman et al., 2017 from a componential perspective).

From the point of view of an EAT, emotional awareness from the response-patterning would require dedicated hardware/software both at the intra-individual (*e.g.*, heart-rate detection, skin conductance, . . .) and inter-individual level (audio/video connection, autonomic emotion recognition from facial expression, . . .). As stated in the introduction, even though these techniques exist and are currently deployed in emotional awareness research in computer-mediated environments, they are not considered for the purpose of the present contribution. It is interesting to point out, though, that a growing body of research considers efferent manifestations as cues of underlying appraisals (Mumenthaler & Sander, 2012; Parkinson, 2013; Scherer & Grandjean, 2007b; Scherer, Mortillaro, Rotondi, Sergi, & Trznadel, 2018). People may therefore infer from efferent manifestations of emotion what kind of evaluation the person is doing of the situation, which then determines the whole emotional episode. This view is consistent with Van Kleef's functional equivalence hypothesis (Van Kleef, 2017) illustrated in the previous chapter. In other words, if people are provided with alternative appraisal cues, the social function of the expression may be considered functionally equivalent.

The action tendency component, on the other hand, is of central concern from an emotional awareness perspective. In fact, according to the CPM, emotion predisposes the organism to act in a certain way, even though the course of action may be blocked or redefined before the concrete execution. Shuman & Scherer (2014) provide some examples with respect to learning:

> For example, feeling curious is associated with approach rather than avoidant behavior. At the same time, as action tendencies are not actions (i.e., overt behavior), one has some flexibility to decide whether or not to carry out the action. This decoupling of stimulus and response distinguishes emotions from reflexes and allows for greater behavioral flexibility. For example, in a classroom, students sometimes need to wait before they can ask their questions. The action tendency associated with curiosity motivates students to ask questions, but it does not dictate that they immediately do so. Similarly, when feeling fear before a test for which one is not prepared, a student needs to suppress the urge to avoid the test and overcome the fear instead by, for example, studying for the test to increase her coping potential.
> — Shuman & Scherer (2014), p. 17

In some important ways, we can derive that emotional awareness is connected to the process of inferring and, if needed, regulate action tendencies in oneself or others. At the same time, the action tendency component has received so far limited attention compared to the other components (Scherer, 2019a, 2022). For instance, Shuman and Scherer (*ibid.*) reckon that epistemic emotions may not be associated to a particular action tendency or adaptive function.

To sum up, even though the Response-Patterning module plays a prominent role in emotion differentiation and has also consequences on behavior especially through action tendencies, its intrinsic nature as well as the need for dedicated hardware and software complicate its transposition in an EAT. As a consequences, elements provided by this module may have to be inferred from contextual or background knowledge by the learner.

3.4.4 The Integration/Categorization Module: Subjective Feeling As Emotion Meaning Making

The integration/categorization module may be considered the *output* of the CPM, because it coalesces information from all the other components into a conscious emotional experience, which can be used to monitor the emotional experience in the person. In fact, the CPM considers the subjective feeling "a holistic cognitive representation that integrates the temporarily coordinated changes of the other components [...], allowing the individual to reach awareness of his/her state and label it – stating that he/she 'has' or 'feels' a particular emotion" (Scherer, Shuman, et al., 2013, p. 281). Contrary to other interpretations of what an emotion is, which requires consciousness for emotion to exist (LeDoux & Brown, 2017; Lieberman, 2019b), according to the CPM the subjective feeling does not necessarily emerge to consciousness, nor fully represents the *true* emotional episode (Scherer & Moors, 2019). On the contrary, as depicted in Figure 3.4, the subjective feeling emerges from the intersection of (A) unconscious, (B) conscious, and (C) semantic/communicative processes.

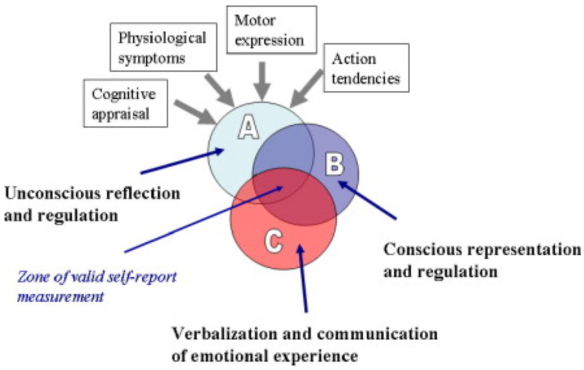


Figure 3.4: The subjective feeling as the central representation of the other CPM's components. The three circles of the Venn diagram show that a valid zone of self-report does not *overlap* with the underlying integration (circle A), nor the intra-personal conscious representation (circle B). Retrieved from Grandjean et al. (2008), Figure 3 in the original article at p. 487.

As the schema illustrates, a virtual *zone of valid self-report* does not completely overlap with the *true* phenomena, meaning that a person can potentially verbalize or communicate (circle C) a subjective feeling that is not representative either or neither of the *true qualia* of circle A or the conscious representation of it represented by circle B. For instance, the person may fail to integrate the whole *signals* from the components in circle A, and therefore not *feel* the episode in spite of the organism experiencing it. This will result in not reporting an emotion. Another possibility is that the person consciously attributes to the *true qualia* a different representation from the *original signal* (circle B), as in the case of interpreting excitement for collaborating with a new colleague as fear. Finally, (circle C) the person may be unable to verbalize her emotion or communicate a different emotion from the one experienced due to confusion or lack of emotion-related vocabulary, as in communicating *frustration* rather than *anger*.

The conscious subjective feeling is therefore tightly related with how the emotion experience is translated into a symbolic representation that can be named, communicated, and operated upon for emotion meaning-making: a lexicalized emotion (Ogarkova, 2013). The combinatorial nature of appraisals in the CPM entails a potential infinite number of emotional experiences. At the same time, it is posited that some combinations are more likely, or more recurrent to occur in everyday life. These patterns – often named *modal emotions* in the CPM – are therefore also more likely to be encoded into the language, for them to be readily accessible in intra-individual or inter-individual symbolic transactions. In this view, commonly adopted emotion words such as *joy*, *anger*, or *surprise* are representative of frequently experienced synchronized patterns. In the meantime, these frequent patterns are nor finite in number, nor forcefully universally shared by all human beings. Due to personal or cultural influences, a pattern can be lexicalized in different manner, or even dispose of a lexical unit in one language but not in another. As a consequence, the conscious subjective feeling in one person is a non-deterministic representation of the underlying *true* episode which is determined at least by (a) the frequency of occurrence of that pattern in the culture/language of reference, and (b) accurate association of that pattern with the frequently adopted lexicalized emotion by the person.

Despite the fact that the integration/categorization module may potentially be misleading with respect to the *true* underlying episode, the CPM subsumes a probabilistic link between the cognitive evaluation of the situation and the resulting conscious subjective feeling. As stated by Scherer (2009b):

> if one knows the results of an individual's event appraisal on major checks, one can approximately predict what kind of emotion he or she will most likely experience (or more precisely, what label the person is likely to use to refer to the experience).
> – Scherer (2009b), p. 1326

In this view, Scherer and Meuleman (2013), used an expert system to measure to what extent it was possible to predict from the CPM's appraisals the subjective feeling experienced by a person. In the test, participants were asked to self-report their appraisals of some situations and provide the subjective feeling(s) they had experienced. Using the self-reported appraisals alone, the system predicted what kind of subjective feeling(s) the participant may have experienced and then matched the results with the one(s) actually reported by participants. Results showed a predictive accuracy by the expert system above 50% for an exact match, and of 90% for a partial correct or appropriate label. An interesting finding of the same contribution concerns participants' need to use, often, more than one subjective feelings to correctly represent their emotional experience, which corroborates how verbalization/communication (circle C) may not completely overlap with the *true* underlying phenomenon.

Further evidence for the central role of appraisal dimensions in the integration/categorization of emotion meaning is provided by studies adopting the GRID instrument (Fontaine et al., 2013). The GRID is a web-based questionnaire that is organized around 142 or 144 emotion characteristics, representing columns of the grid, each one pertaining to one of the five CPM's components. Each row can be represented by one lexicalized emotion. Participants can thus rate this semantic word according to how likely a person in their culture is to manifest each of the characteristics of the emotion. The *profile* that emerges from the combination of rows and columns can then be adopted in different analyses and for different purposes (Fontaine et al., 2007, 2007; Gentsch et al., 2017; Gillioz et al., 2016; Scherer & Fontaine, 2018).

Fontaine and colleagues (2007), for instance, adopted the GRID instrument in a cross-cultural studies using 24 prototypical emotion words in 3 languages. Using a principal component analysis (PCA), the author found evidence for the need of 4 underlying dimensions to account for similarities and differences between emotion words. The authors identified these dimensions as evaluation-pleasantness, potency-control, activation-arousal, and unpredictability. A similar, but more comprehensive study, conducted this time in over 20 languages and countries yielded corroborating results (Fontaine et al., 2013). The authors found evidence for a four-dimensional structure, with the four factors identified as *Valence*, *Power*, *Arousal*, and *Novelty*. Gillioz and colleagues (2016) extended the analysis to 80 emotion terms and confirmed the emergence of a four-dimensional structure with the same aforementioned factors. Incidentally, one of the advantages of the GRID instrument is that the PCA projects the emotion words into an affective space characterized by the underlying dimensions, a feature that will be of central importance in the EAT proposed in this contribution.

Scherer and Fontaine (2018) performed a secondary analysis on the large datasets from over 20 languages and countries (Fontaine et al., 2013) with the specif intention to test whether the feeling component factors may be explained by the appraisal components, and whether the response-patterning factors played a mediating function. The author found evidence that the appraisal components explained the results of the subjective feeling components, and that the effect was almost fully mediated by the response-patterning components. In other words, when one knows the appraisal profile, the response-patterning adds little value to the explanation.

Finally, Gentsch and colleagues (2017) adopted the GRID and a modified version, the Achievement Emotion CoreGRID, to test whether the semantic structure of emotion words may be influenced by contextual factors, for instance a general, non-contextual situation vs. an achievement situation characterized by either failure or success. The authors suggest that their analysis confirms the emergence of a four-dimensional structure with the same factors (*Valence*, *Power*, *Arousal* and *Novelty*), which is consistent regardless of the general or achievement context. On the other hand, specific features of emotion were rated differently depending on the context. More specifically, the appraisal components ratings were the ones more likely to be affected by the context, which is aligned with appraisal theories' central tenets. It is worth noting, tough, that Gentsch and colleagues (*ibid.*) adopted a MANOVA followed by post-hoc uni-dimensional t-test according to the significance of the multivariate test, a practice whose validity is subject to critics (Warne, 2014).

To sum up, there is consistent evidence of a direct link between appraisal criteria and the holistic, conscious subjective feeling, which integrate the whole phenomenon. This link is particularly evident when the subjective feeling is represented by a lexicalized emotion: a word carrying semantic emotional meaning. The link is nevertheless probabilistic in nature, and can be influenced by contextual as well as individual characteristics such as emotion knowledge, emotional vocabulary, or culture. At the same time, lexicalized emotion can also be projected into affective spaces determined by dimensions driven by a componential approach. In this regard, there is consistent evidence that four dimensions are necessary to account for the differentiation of emotional experience.

3.5 Emotional Competence from a Componential Perspective

The adoption of the CPM as the framework of reference presents another interesting aspects: disposing of an underlying mechanism upon which the overarching concept of *emotional expertise* as defined by Hoemann et al. (2021) can be clarified. In fact, Scherer (Scherer, 2007) proposes to adopt the term *emotional competence* and postulates that this phenomena can be informed by the CPM.

Scherer defines emotional competence as “the capacity that is at issue, varying over individuals, as differential degrees of *competence in using the emotion mechanism* as it has been shaped by evolution” (Scherer, 2007, p. 102, italics in the text). The term competence is considered at an intermediate level in a continuum represented by relatively stable traits, such as abilities, on the one side, and more specific and contextualized skills on the other side. This conceptualization is consistent with other scholars considering advantages inherited by some form of emotional know-how as malleable, especially through training (*e.g.*, Mayer et al., 2016). Emotional competence can be divided in two major domains: emotion production and emotion perception. The two are then further characterized by three sub-competences: 1) the appraisal competence, (2) the communication competence, and (3) the regulation competence.

3.5.1 Appraisal Competence

The appraisal competence consists in emotion elicitation and emotion differentiation (Scherer & Moors, 2019). On the one hand, a competent emotional agent is able to detect features of the environment that may benefit from an emotional response. For a student, for instance, it is important to rapidly detect events that are central to the learning process such as novel and relevant information; interactions with colleagues that may lead to unpleasant consequences; concepts that exceeds one’s coping potential in processing them; or behaviors that violates norms of conduct such as plagiarism.

The detection of emotion-eliciting events is nevertheless not enough to unfold the *most adequate* response. Emotional differentiation consists thus in the generation of a *valid* appraisal profile, where valid stands for a profile that is most likely to generate an adaptive emotion. For instance, detecting that the content of a lesson is exceeding one's coping potential may lead to confusion or boredom, with the former most likely to enhance cognitive re-equilibrium, whereas the latter may trigger abandoning altogether (R. S. J. D. Baker et al., 2010; D'Mello et al., 2014). Emotion differentiation is tightly related with the idea that even if there is no emotion that is intrinsically positive or negative for learning, there is an emotion that is more instrumental to learning processes or outcomes at a given moment, with a given appraisal profile.

At the same time, there may be learners that are more or less inclined to experience some emotions more frequently than others. Scherer (2021a), for instance, has gathered empirical evidence suggesting that an *appraisal bias* – that is, a stable tendency, often dysfunctional or unrealistic, to evaluate events in similar ways – may be at the core of emotion dispositions. The intervening role of appraisal criteria may in this sense provide a more fine-grained causal mechanisms compared to the more unspecified idea of affective traits (*ibid.*).

An EAT may thus sustain or elicit appropriate appraisal of a situation and emotional differentiation by providing learners with a set of criteria against which events may be evaluated. The appropriate emotional response in learning settings, though, may be more difficult to establish compared to situations in which the well-being of the organism is more directly implicated. According to Shuman and Scherer (2017, p. 14) "emotions can be good or bad for learning depending on how similar requirements posed to us in modern situations are to demands for survival in the past". This kind of similarity may require a more abstract, meta-cognitive and conceptual evaluation of the situation at hand, and an EAT may in this sense be adaptable to appraisals that are more likely to occur given the specifics of the learning activity.

3.5.2 Communication Competence

The communication competence is two-fold. On the production side, it concerns the expression of a strategic signal that maximizes correct interpretation from others, minimizing ambiguous inferences. The signal may be verbal or non-verbal, but in both cases, it is in an emotionally competent agent's interest to make others aware of her emotional episode, for it has been triggered by an important event. For instance, it is in a student's interest to express her *contempt* towards a free-rider, for this behavior to cease; or *gratitude* towards an helping colleague, for this behavior to repeat in case of need.

On the perception side, the communication competence extends on inter-individual settings, by requiring accurate decoding of the strategic signal emitted by another person. As posited by the Emotion As Social Information framework (Van Kleef, 2010, 2018), accurate interpretation of emotional signals are valuable information both for inferring causes and consequences of behavior in others, but also for the person herself.

An EAT may therefore play a prominent role in facilitating the communication of the most accurate strategic signal from the emitter to the decoder. At the same time, it may also train learners to improve their emotion communication competence, for instance by widening the *toolbox* of strategic signals that may be sent or received (Erbas et al., 2015).

3.5.3 Regulation Competence

From an appraisal driven perspective, emotion regulation may be seen as an additional evaluation layer, which for once does not evaluate the situation *per se*, but rather the emotional response that has been produced. According to Scherer (2007), one of the central functions of emotion regulation is to avoid the persistence of emotion that have been triggered by inappropriate evaluation of the situation. This mechanism is particularly suited to appraisal theories, which advocates emotion as an ongoing process that may be refined over iterations in order to produce the *optimal* emotional response.

As seen in Section 1.1.3, regulatory processes are central tenets of socio-emotional competences (Brackett et al., 2019; Hoffmann et al., 2020) as well as the regulation, co-regulation and socially shared regulation of learning (Järvelä et al., 2016; Miller & Hadwin, 2015). An EAT may therefore enhance regulation processes by providing information about how a situation has been appraised, augmenting the chances that both the learner herself and colleagues may assess to what extent the evaluation of the situation is congruent with the learning dynamics at stake.

3.6 Summary

This chapter focused on the *emotion* noun in the Emotion Awarenss Tool acronym, in an attempt to define what may be the *unit* of information in emotional awareness. The chapter first highlighted how difficult, but central, is the definition of *what an emotion is.* The choice of appraisal theories of emotion and, more specifically, the Component Process Model as the framework of reference for this definition has been motivated through the integral role of the cognitive evaluation against a set of criteria, which are of major concern for the person. This evaluation determines: (1) what is the kind and intensity of the emotional experience elicited; (2) what is the most likely symbolic reference the person may associate to the episode for individual meaning making or interpersonal strategic signalling; and (3) whether the elicited emotion is *appropriate* given the situation at hand, that is, if it has been assessed *competently* or some form of individual or social regulation is needed. The link between appraisal dimensions, represented by a number of dimensional criteria, and the holistic subjective feeling, which may be symbolized by a lexicalized emotion, is therefore retained as the emotion structure to be implemented into an EAT. It is posited that the functions of emotional awareness in computer-mediated learning environments can best be served by building an EAT upon the causal mechanism underlying emotion elicitation, differentiation, communication and regulation provided by the Component Process Model framework. If this mechanism can be even partially solicited by the EAT, it is argued, learners can extrapolate instrumental meaning-making from emotional awareness, which can be integrated in the self-regulation, co-regulation or socially shared regulation of learning.

This chapter also concludes the first part of the thesis, in which the theoretical foundations of the whole contribution have been set forth. The upcoming Part II will therefore focus on the implementation of an EAT that transposes this theoretical background into concrete elements of a Graphical User Interface: the proof of concept.

Part II

Proof of Concept

Chapter 4

Previous Work on a Prototype: the Dynamic Emotion Wheel

This chapter resumes previous work that I have conducted in my Master thesis[1] on the subject of Emotion Awareness Tools (Fritz, 2015). The Master thesis was a *spin-off* of the Emotion Awareness Tools for Computer-Mediated Interactions (EATMINT) project and was first mainly focused on a user experience perspective, lacking therefore most of the theoretical background set forth in Part I. The Master thesis nevertheless lead to a conceptual prototype, which I named Dynamic Emotion Wheel (DEW). The prototype was first presented at the International Society for Research on Emotion (ISRE) conference (Fritz et al., 2015), and later in education technology conferences (Fritz, 2016a, 2016b). The interest elicited in both affective sciences and education technology communities corroborated further conceptual and technical work on the prototype. The chapter starts by briefly outlining the aim and scope of the EATMINT project as the context through which a series of initial requirements for an EAT were set forth. Second, it illustrates the Geneva Emotion Wheel (Scherer, 2005; Scherer, Shuman, et al., 2013), a self-report tool based on the Component Process Model depicted in Section 3.4. The Geneva Emotion Wheel served as inspiration to produce the core principle of the Dynamic Emotion Wheel, illustrated in the third section of the chapter. The section presents an overview of the prototype, an affective space derived from previous work in the EATMINT project, and also the main results of a usability test conducted in experimental settings. Finally, the theoretical and technical extensions to pass from a prototype to a proof of concept are outlined.

4.1 The EATMINT Project

The Emotion Awareness Tools for Computer-Mediated Interactions (EATMINT) project was conducted as part of the National Center of Competence in Research (NCCR) *Affective Sciences* by a group of scholars integrating affective sciences, computer science, education

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

technology, education and learning sciences (Cereghetti et al., 2015; Chanel et al., 2016, 2013; Gaëlle Molinari, Bozelle, et al., 2013; Gaëlle Molinari, Chanel, et al., 2013). When the project started in mid 2011, the topic of emotional awareness in computer-mediated collaboration was almost uncharted (Eligio et al., 2012; Feidakis et al., 2011). The team therefore initiated the project to explore two intertwined objectives:

1. Designing Emotion Awareness Tools, which allow either the explicit or the autonomic sharing of emotions in computer-mediated interaction;
2. Studying the impact of emotional awareness on collaboration both at intra- and inter-personal levels, with respect to the emergence of strategies that improve collaboration such as monitoring, regulating and reflecting on emotions.

A first step in the project consisted in the set up of exploratory experiment, from which different types of data could be collected according to complementary research perspectives. First, to what extent emotional awareness influences perceived emotions after collaboration and the perceived quality of interaction (Gaëlle Molinari, Chanel, et al., 2013). This part of the study has been thoroughly illustrated in the related works of Section 2.3 about the inter-personal function of emotional awareness. Second, to what extent the perceived interaction during a collaborative task is linked to users' emotional traits, and how an emotional awareness tool can influence this relationship (Gaëlle Molinari, Bozelle, et al., 2013). Third, to what extent physiological and eye-movement coupling predict situations that may benefit from emotional awareness during the collaborative process (Chanel et al., 2013).

The experimental setup was the same already described in Section 2.4 about emotion in computer-mediated learning environments. In this regard, the EATMINT project members reckoned that the EAT used in pilot experiments – already depicted in Figure 2.4 and repeated hereafter in Figure 4.1 – showed some limitations, and therefore planned to implement a second version of the tool to be used in future studies. I was involved in this process first as an internship in the project, and later as the subject of my Master thesis (Fritz, 2015). This section starts with a more comprehensive description of the tool. Then, it illustrates a series of requirements for the next version of the EAT, which were discussed with the EATMINT project's team. Finally, it gives a brief overview of the interaction design method which was applied to meet these requirements.

4.1.1 Description of the EATMINT's First Version of an EAT

The tool in the pilot studies of the EATMINT project was persistently located at the right edge of the screen, occupying the whole height and about a fifth of the width. The tool was vertically divided in two areas of almost the same size.

The lower part of the tool was dedicated to the expressing-displaying function of an awareness tool. Users could encode their emotional experience by pressing one of 20 buttons, each labelled with a lexicalized emotion in the form of an adjective (e.g. *attentive*, *irritated*, *relaxed*, etc.). The list of discrete emotions was defined by combining a previous study about affective dynamics during complex learning (D'Mello & Graesser, 2012) and two pilot studies that aimed at identifying the most frequent and intense emotions felt during a situation of

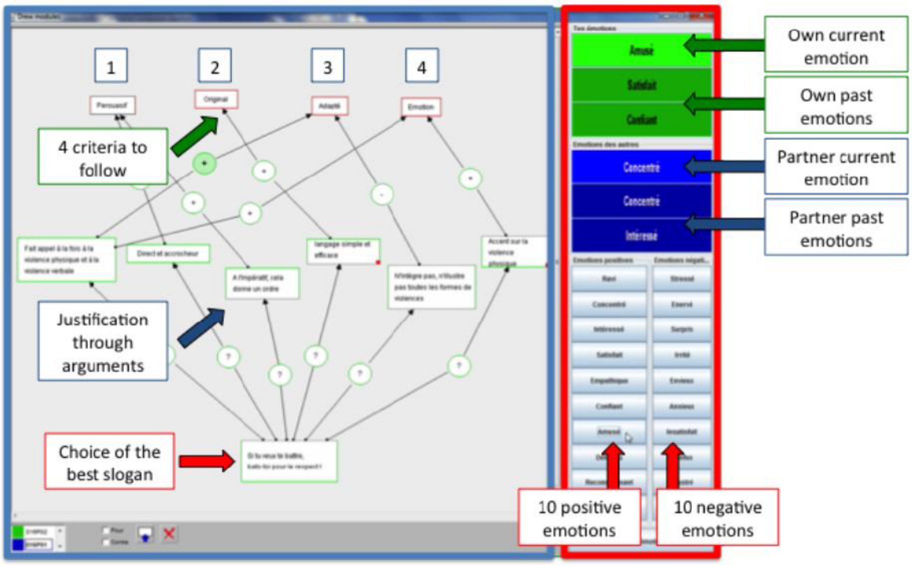


Figure 4.1: The argument graphic tool and the EAT adopted in Gaëlle Molinari, Chanel, et al. (2013), from Figure 1 in the original article (repeated from Figure 2.4).

collaboration, real or imagined. The 20 discrete emotions were equally divided in *positive* and *negative* labelled groups, available side by side, but did not follow any particular order within each list (see Table 4.1). At the bottom-end of the screen, a "no-emotion" button spanning across the two columns was also available.

The upper side of the tool was primarily concerned with the perceiving-monitoring function of an awareness tool. Participants could in fact see the last three emotions they had expressed and the last three emotions manifested by the partner. Their own emotions

Table 4.1: List of the 20 emotion adjectives used as buttons in the EAT. Our translation from the original French word in parentheses is based on a document by the Geneva Emotion Research Group (1988). The order in the table reflects the order of the buttons in the interface.

	Positive emotions	Negative emotions
1	Delighted (Ravi)	Stressed (Stréssé)
2	Attentive (Concentré)	Annoyed (Enervé)
3	Interested (Intéressé)	Surprised (Surpris)
4	Satisfied (Satisfait)	Irritated (Irrité)
5	Empathetic (Empathique)	Envious (Envieux)
6	Confident (Confiant)	Anxious (Anxieux)
7	Amused (Amusé)	Disappointed (Insatisfait)
8	Relaxed (Détendu)	Confused (Confus)
9	Grateful (Reconnaissant)	Frustrated (Frustré)
10	Relieved (Soulagé)	Bored (Lassé)

appeared on top, as a vertical list of text-centered cases with a green background. The very last emotion had a lighter background and resulted first in the list. This case also served as a text input where participants could type an emotion not in the list, having thus also an expressing-displaying function. Their partner's emotions appeared just below, in exactly the same way, except for a blue background and the first light-blue case that was not editable.

As a result, each participant could voluntarily express their emotional experience in one of three ways: (1) by clicking one of the 20 emotion labelled buttons; (2) by typing an emotion not in the list directly in the first light-green case; or (3) by clicking the "no-emotion" button at the bottom-end of the screen. The expressed emotion would then appear in the upper light-green case of the participant's own screen, and, respectively, in the upper light-blue case of the partner's screen.

4.1.2 Requirements for a Second Version of an EAT

Following the analysis of the experiment from the team and the feedback provided by participants – who found, for instance, the unordered disposition of the discrete emotions labels not helping – a requirement analysis was conducted with the members of the EATMINT project. At first, the requirements were set forth from a general, brainstorming perspective and it was the aim of the internship to try to put them together in a coherent interface. The requirements analysis identified five main characteristics of the tool.

First, it should keep the voluntary self-report approach. Even though the project's aim contemplated both autonomic and voluntary tools, this specific EAT was intended to follow-up on the voluntary disclosure of the emotional experience.

Second, the tool should be inspired by an emotion theory. One of the considered candidate was Pekrun's Control-Value theory of achievement emotions (Pekrun, 2006; Pekrun & Perry, 2014), already cited in Section 1.1.1 and 3.3. Given that the theory is deeply rooted in the appraisal family, it was also suggested that the tool should integrate a dimensional approach representing the Control and Value appraisals rather than relying exclusively on discrete emotions.

Third, an additional reason to switch to a dimensional approach concerned the reduction of the number of discrete emotions on screen, which was perceived as a source of cognitive overload from participants. The sheer number of discrete labels available, combined with the lack of a specific order, forced participants to perform a seek-and-evaluate task (Rouet & Tricot, 1995) every time they had to click on a proposed label.

Fourth, the presence of 20 buttons on the interface also reduced the space available for the perceiving-monitoring function of the tool. Only the last three emotions for each participant appeared on the interface without any further information such as the time when they occurred or in which order they appeared considering the dyad as a whole. In other words, it was possible to establish in which order each participant had expressed their last emotions, but not if the very last emotion of the dyad belonged to the first or second participant. It was also impossible to tell, from the interface alone, if the order of the emotion followed a one-each turn, or a participant expressed three or twenty emotions in a row. Consequently, for the new version of the tool, the EATMINT project members wanted to enhance the possibility for participants to (1) dispose of the evolution of their own emotional state over time, and (2) compare this evolution with their partner's one.

A last requirement, of a more general scope, concerned the flexibility of the tool with respect to theoretical and technical features. On the theoretical level, the new version of the tool should provide experiments with as much leverage as possible about the underlying theoretical framework, for example the number and labels of the emotions. On the technical level, the interface of the tool could also be considered as a potential factor in the experimental settings. For example, a group of participants could dispose of an interface that particularly highlights inter-personal comparison, whereas the other group an interface that provides more insight on the intra-personal evolution over time. Finally, the tool should also provide a way to set up emotional awareness in experimental settings without specific technical knowledge, and dispose in the meantime of collected data in a practical form for exportation and subsequent analysis.

4.1.3 Interaction Design Process

The establishment of requirements was the first step in an interaction design and user-centered approach (Cooper et al., 2007; Lallemand & Gronier, 2017; Rogers et al., 2011), which is driven by the underlying assumption to take the end-user perspective from the very onset of the process. The other steps advocated by this iterative method are: the design of alternatives, often based on an analysis of existing and concurrent products; building prototypes at various levels of fidelity; and finally evaluating to what extent the prototype meets users' functional needs and perceptual preferences. The full process that eventually lead to the Dynamic Emotion Wheel is thoroughly depicted in my Master thesis (Fritz, 2015). Hereafter, only the most important steps in the process are illustrated, starting from the identification of the closest existing product that presented some of the features aimed by the new version of the EAT.

4.2 The Geneva Emotion Wheel

The Geneva Emotion Wheel (GEW) is an emotion self-report tool based upon the Component Process Model (CPM) theoretical framework illustrated in Section 3.4. At the time of writing, the tool is in its third version, having received modifications over the years (Scherer, 2005; Scherer, Shuman, et al., 2013; Shuman & Scherer, 2014). This section starts with a description of the tool and how it relates to the CPM. It then illustrates an empirical use of the tool, which resonates with emotional awareness. Finally, it assesses the limitations of the tool with respect to the features retained as a central tenet for the implementation of an EAT in the current contribution.

4.2.1 Description of the Geneva Emotion Wheel

The GEW was initially created under the assumption that "it might be worth investing in the development of an instrument capable of combining the advantages of the precise differentiation provided by natural language labels with the simple organizational structure afforded by a two-dimensional space" (Scherer, Shuman, et al., 2013, p. 283), see also Section 2.4. The assumption also posits that being emotion determined by the cognitive evaluation of the situation described in Section 3.4.2, the two-dimensional space of the GEW should be determined by appraisal dimensions (Scherer et al., 2006). In the meantime, the discrete emotions – represented by lexicalized emotions as manifestations of the subjective feeling illustrated in Section 3.4.4 – should be arranged according to the affective space that emerges from the combination of the two underlying dimensions (Scherer, 2005; Scherer et al., 2006; Scherer, Shuman, et al., 2013). The GEW, depicted in its third version in Figure 4.2 (Scherer, Shuman, et al., 2013), thus, places 20 lexicalized emotions in the form of natural language words around a circumplex determined by the *Valence* dimension on the horizontal axis, and the *Control/Power* dimension on the vertical axis.

The choice of the two dimensions is warranted by the fact that there is a clear correspondence between the two dimensions and the sequential evaluation checks of the CPM (see Table 3.1. The widely adopted dimension of *Valence* corresponds to the stimulus evaluation checks pertaining to the *Relevance* and *Implications/Consequences* groups, namely intrinsic pleasantness and goal conduciveness. The same applies to the stimulus evaluation check of the *Coping potential* group, which can be related to the dimension of *Control/Power,* often adopted in dimensional approaches (Bradley & Lang, 1994; Broekens & Brinkman, 2013; Osgood, 1952). Previous studies on appraisals criteria have also confirmed that the two appraisal groups are also the ones that contribute the most to emotion differentiation (Scherer, Shuman, et al., 2013), even if they are not enough, as previously mentioned, to account for all the gamut of experiences (Fontaine et al., 2007; Gillioz et al., 2016). It is also worth noting that the authors discarded the most commonly adopted dimension of *Arousal/Activation* (Brackett et al., 2019; Feidakis, 2016; Feidakis et al., 2011; James A. Russell, 1980). The main reason behind this choice concerns the difficulty in distinguishing between the intensity of the subjective feeling from the intensity of the bodily excitation (Scherer, Shuman, et al., 2013). Furthermore, the authors also point out that:

> while most lay persons have little problem evaluating the positivity or negativity of a feeling (or event) and the approximate degree of their felt arousal, the resulting point in two-dimensional space has no specific meaning for them and cannot be communicated to others. It would seem very strange to tell someone that I feel 2.3 positive and 1.6 aroused.
> — Scherer, Shuman, et al. (2013), p. 283

The 20 lexicalized emotions are thus positioned according to a value along the *Valence* and the *Control/Power* dimensions. As a consequence, each discrete emotion belongs to one of the four quadrants of the circumplex, determined by the combination of positive or negative *Valence*, and positive or negative *Control/Power*. The lexicalized emotions of version 3.0 are in clockwise order: *interest*, *amusement*, *pride*, *joy* and *pleasure* for the positive *Valence* and high/positive *Control/Power* quadrant; *contentment*, *admiration*, *love*, *relief* and *compassion* for the positive *Valence* and low/negative *Control/Power* quadrant; *sadness*, *guilt*, *regret*, *shame* and *disappointment* for the negative *Valence* and low/negative *Control/Power* quadrant; and *fear*, *disgust*, *contempt*, *hate*, and *anger* for the negative *Valence* and high/positive *Control/Power* quadrant.

Furthermore, from the origin of the axes stem 5 circles for each of the lexicalized emotions. The circles grow in size as they get closer to the edge of the circumplex, so that the first circle is the smallest, and the last one the biggest. Participants can therefore choose, first, the *row* of circles that corresponds to the subjective feeling they are experiencing and, second, the intensity of the emotion as a function of the circle's size. At the same time, respondents are also provided with the "None" and "Other" options available in the center of the circumplex, which is considered a good practice in emotion self-report (Mortillaro & Mehu, 2015).

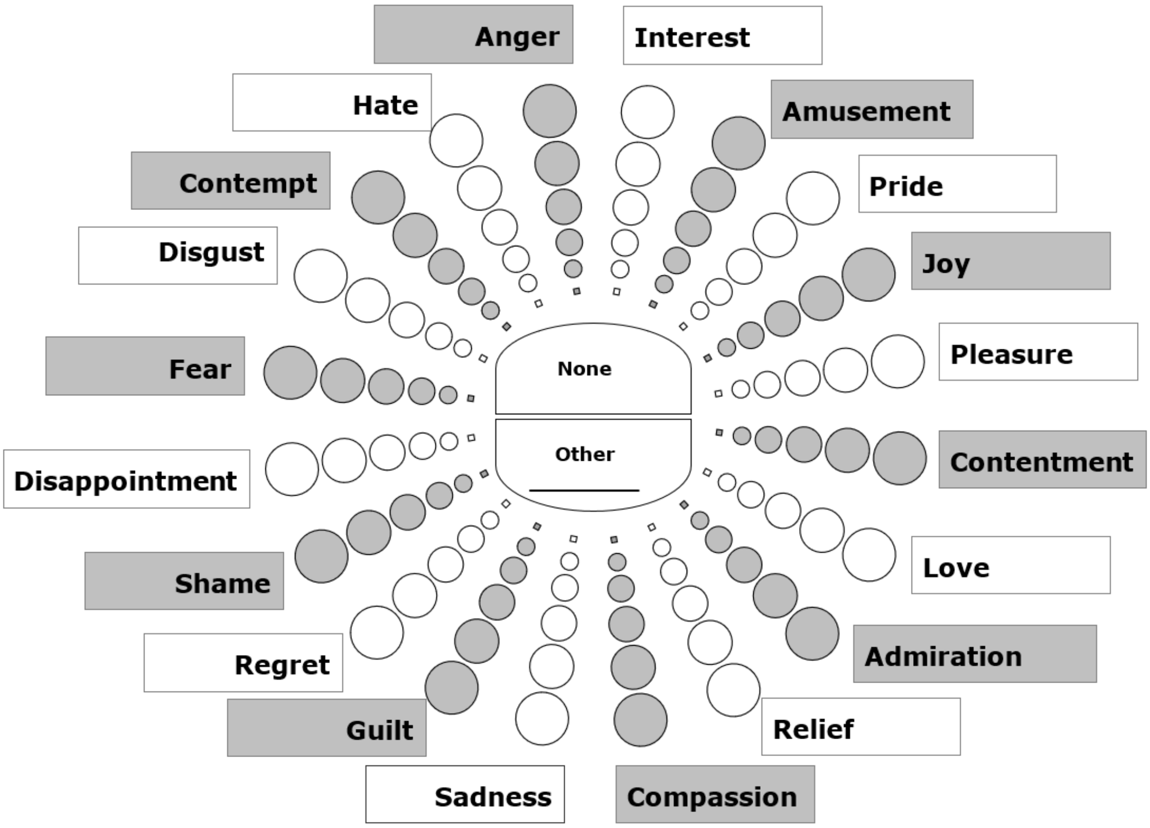


Figure 4.2: The third version of the Geneva Emotion Wheel (Scherer, Shuman, et al., 2013)

It is worth noting that, whereas the two dimensions have been constant over the evolution of the tool, the number and labels of the lexicalized emotions have been modified with each of the three versions, even though the aim has been consistently to provide users with intuitive, widely adopted natural language words representative of major emotion families (i.e. modal emotions). In fact, the positioning of the discrete emotions proved to be particularly difficult. If the position of the lexicalized emotions for version 1.0 and 2.0 was empirically validated for most of them on the *Valence* dimension, the same was not true for the *Control/Power* dimension (Sacharin, Schlegel, & Scherer, 2012). This aspect will be further developed in Section 4.2.3. For this descriptive passage, it is worth mentioning that the disposition of the 20 lexicalized emotions of the third version of the GEW has been determined using the GRID instrument mentioned in Section 3.4.4. The authors who have been working on the third version of the GEW also suggest that a custom choice of lexicalized emotions is possible, but should ideally be determined by the GRID instrument. In this regard, the MiniGRID or the CoreGRID (Scherer, Fontaine, & Soriano, 2013), which are shorter version of the GRID, can be adopted to limit the number of ratings needed to determine the position of each lexicalized emotion.

To sum up, the design of the GEW provides different interesting features from the point of view of an EAT. First, it implements an emotion structure in a self-report tool by combining appraisal dimensions and subjective feelings into a theoretically and empirically driven underlying structure. Second, the subjective feeling is represented by lexicalized emotions that are meant to be widely accessible in the emotional vocabulary due to their frequency in everyday language. Third, it allows the assessment not only of the kind of emotional experience, but also its intensity through the size of the circles. Fourth, the whole is *packed* into a user-friendly interface that can be immediately understood by respondents and produce reliable emotion self-report (Caicedo & van Beuzekom, 2006).

4.2.2 The Geneva Emotion Wheel as Emotional Awareness Tool

The GEW has been adopted in a wide variety of empirical contributions ranging from consumer attitudes towards a product to the use of the tool from an experience sample method perspective (Scherer, Shuman, et al., 2013). Even though the primary aim of the GEW is the accuracy in individual self-report of the subjective emotional experience, there has also been attempts to use it at the inter-personal level.

In particular, Veronique Tran et al. (2012) adopted the GEW – in what seems version 1.0, which is probably why the version is unspecified in the article – to investigate the role of emotions in decision-making at the collective level in management teams. Groups of 4 to 7 participants were formed to simulate the collective conduction of a company during fictitious long-term periods of 8 years, represented by *stints* of three hours in the study. Participants expressed on the GEW, using a pencil and paper version of the tool, two emotions at the beginning, at the mid-point, and at the end of each period for a total of 6 emotions. The emotions appeared on the same sheet, and each emotion was identified by a code composed by the letters B (beginning), M (middle) or E (end), as well as a sequential number from 1 to 6. The code of the emotion was placed in one of the circles of the interface representing the subjective feeling and its intensity. The individual rating of the emotional experience was then discussed with the other members of the team, who negotiated to define the two emotions the team perceived as prevailing during the decision process and reported them as *consensus emotions* on another sheet with the GEW. This fact created the conditions for emotional awareness to emerge in the teams. Among the practical implications of the study, in fact, the authors highlight what follows.

> Over the course of this study, it could be observed that participants often used the Emotion Wheel as a medium to discuss their emotions freely with their colleagues and it became part of the norms of the teams to do so. Thus, discussing emotions yields self-awareness and awareness at the group level when, for example, participants discuss their group consensus emotion. In addition,

> by mapping emotions on the [Geneva] Emotion Wheel on a regular basis, everyone can see the evolution of the emotional climate and team members can proactively manage it.
> — Veronique Tran et al. (2012), p. 22

Having their emotions expressed on the underlying structure represented by appraisal dimensions, in fact, allowed participants to notice not only the switch from a discrete emotion to another, but also the switch from a position in the circumplex to another, for instance a change from a quadrant to another. To sum up, the implementation of a theoretically driven emotion structure facilitated the emergence of emotional awareness, with the consistency and familiarity of the structure becoming a pivotal element in the consideration of emotions within the task at hand. This empirically corroborates the potential usefulness in implementing an emotion structure into an EAT.

4.2.3 Limitations and Missing Features in the Geneva Emotion Wheel

Notwithstanding the many features of the GEW that corroborate the intrinsic quality as a self-report instrument, as well as the possibility to adopt it in circumstances linked to emotional awareness, the tool presents a number of limitations and missing features, particularly from the point of view of an EAT with the characteristics sought after in the present contribution. Most of them are from a technical or user-experience perspective. Others, though, are grounded more at a conceptual level that interacts also with the theoretical underpinning of the tool. These are headed first, leaving more practical considerations as a follow-up.

As mentioned in the description of the GEW, the two underlying dimensions from which the emotional structure is denoted are *Valence* and *Control/Power*. In an attempt to empirically validate the disposition of the three versions, the *Control/Power* dimension was hard to match (Sacharin et al., 2012; Scherer & Fontaine, 2018; Scherer, Shuman, et al., 2013). The problem was particularly exacerbated when the disposition of the lexicalized emotion was determined through direct rating on the two dimensions. In one of this attempt, for instance, the *Valence* dimension was depicted as “the situation is experienced as (un)pleasant and enjoyable (disagreeable) and/or is likely to have positive and desired (negative and undesired) consequences for the person”. *Control/Power* was in turn stated as “the person believes that he/she can (cannot) influence the situation to maintain or improve it (if desired)” (Sacharin et al., 2012). Results highlighted a substantial overlapping, or correlation, between the *Valence* and the *Control/Power* dimensions, especially with the top-right (negative *Valence* and high/positive *Control/Power*) and bottom-left (positive *Valence* and low/negative *Control/Power*) quadrants. As a consequence, the perception of the two dimensions is not orthogonal as the underlying structure may imply. Another problem was also encountered with the GRID instrument, though, for the features related to the *Control/Power* dimension did not yield the theoretical importance associated with the Coping potential in emotion elicitation and differentiation (Scherer & Fontaine, 2018).

Two main reasons are posited for these difficulties (Sacharin et al., 2012; Scherer & Fontaine, 2018; Scherer, Shuman, et al., 2013). First, the notion of *Control* may be too abstract for non-psychologists. As a result, even the sheer way by which the dimension is instantiated through wording may be misleading. It may thus be a matter of trial and error before a viable description is found in order to align the theoretical and the perceived character of the dimension. To this extent, the CoreGRID already modified some wording in the rating of related features (Scherer, Fontaine, et al., 2013). Second, the *Control/Power* dimension may be intrinsically *valenced* (Shuman et al., 2013). Situations with low control are perceived as generally unpleasant, whereas situations with high control are perceived as pleasant. As summarized by Scherer and Fontaine:

> It is exceedingly difficult to construct items that allow one to obtain valid assessments of control/power/coping appraisals, partly because of the strong relationship to valence (it is good to have high power).
> — Scherer & Fontaine (2018), p. 9

Another limitation linked to the underlying appraisal dimensions is that, in the GEW, these dimensions are not manifest. The problem with the *Control/Power* dimension surfaces only from a retro-engineering mechanism, but on the tool itself, there is no explicit mention of neither *Valence* nor *Control/Power*. The two appraisal dimensions are used to organize the lexicalized emotions, but there is no explicit guarantee that a person choosing *shame* (bottom-left quadrant) is *aware* of having evaluated the situation as unpleasant and with little coping potential – that is, unless the respondent is informed beforehand of the semantic value of the interface. Without this kind of training, even the emotional awareness role played by the GEW in Veronique Tran et al. (2012) may be reassessed. According to more thorough information about the procedure available in Véronique Tran (2004), participants received definitions of the emotions on the GEW taken from a dictionary. As a consequence, there was probably some information about *Valence*-like features, but less likely about an underlying theoretical construct such as *Control/Power*. When tracing the evolution of the emotional experience between the quadrants, participants may have not been completely aware of the *dimensional* leaps undertaken, if not by the inference that adjacent emotions on the interface were more akin than those afar.

From a more practical standpoint, the GEW is also doomed to occupy a considerable share of a screen, as it is the case for the *emot-control* (Feidakis et al., 2014; Feidakis et al., 2013) tool illustrated in the related works of Section 2.4. The strength of this kind of interfaces is to provide an organized structure to the discrete emotions, for respondents to dispose at the same time of all the available options. As a consequence, this design is not very responsive to scale in proportion to the number of available options. A greater number of discrete emotions would shrink the interface to the point of overlapping boundaries, especially for the circles to rate the intensity. A lower number of options, without resizing the whole interface, would tear options far apart, injecting negative space in between, which would be of limited visuo-spatial use (Hegarty, 2011).

Another shortcoming of the GEW is its potential to fulfill the perceiving-monitoring function of an awareness tool. The size of the interface leaves little space to provide a dedicated area for contemplating the emotion expressed through the tool, especially from a longitudinal perspective. As in Veronique Tran et al. (2012), it would be possible to mark directly on the circles the expressed emotions with some form of sequential encoding to emulate time, but this mechanism may easily overcrowd the interface after a few responses. One may duplicate the interface, for instance leaving the upper GEW for expressing-displaying, and the bottom GEW for perceiving-monitoring; but the tool would in this case occupy twice the already considerable size of one GEW alone.

To sum up, The Geneva Emotion Wheel presents a number of theoretical and empirically validated features that are appealing even for an EAT. First, it is based on self-report, which entails the conscious, conceptual effort to investigate one's own emotional experience. Second, this introspection is guided by the implementation of an emotion structure combining appraisal dimensions and subjective feelings. Third, the structure also forms an affective space, which can be adapted by using appropriate steps in validating the characteristics of an alternative disposition if desired. At the same time, the tool also presents some technical shortcomings and missing features that limit the possibility to be adopted *as-is* to perform both the expressing-displaying and the perceiving-monitoring functions of an EAT. As a consequence, it was considered worth investigating the development of a new tool.

4.3 The Dynamic Emotion Wheel

The Dynamic Emotion Wheel (DEW) was originally conceived as a web-based application targeting the EATMINT project's requirements outlined in Section 4.1.2. It was therefore primarily driven by an inter-personal perspective about emotional awareness consistent with the EATMINT interest in computer-mediated collaboration (Fritz, 2015, 2016a, 2016b; Fritz & Bétrancourt, 2017; Fritz et al., 2015). As the name implies, the tool is deeply inspired by the Geneva Emotion Wheel (Scherer, 2005; Scherer, Shuman, et al., 2013; Shuman & Scherer, 2014) described previously. This section starts by laying out the core principle of the DEW – consisting in the *dynamic* element of the acronym – which represents the main conceptual and usability difference with the Geneva Emotion Wheel. Second, it describes how this core principle determine the whole interface of the tool. Next, it illustrates how the principle can be adapted to a different affective space by introducing the EATMINT circumplex (Fritz & Bétrancourt, 2017). The sections ends with the results of a usability test conducted in experimental conditions as the last step of a first iterative process in the interaction design and user-centered method adopted to create and later incrementally improve the tool (Fritz, 2015).

4.3.1 The Core Principle of the Dynamic Emotion Wheel

The core principle of the DEW is rather straightforward: in a self-report condition, emotional awareness is determined by how the *encoding* of emotion is prompted. Everything follows from and depend on this process. As a consequence, the mechanism by which emotion is encoded should adhere as much as possible to the mechanism by which emotion is elicited, differentiated and symbolically represented in oneself and in communicating with others. This principle is implemented by the DEW with three sequential and coordinated steps, illustrated graphically by the low-fidelity wireframe in Figure 4.3.

The first two steps are the same type of action and are driven by the following assumption: rather than using appraisal dimensions to organize the lexicalized emotions as in the Geneva Emotion Wheel, the very same appraisal dimensions are used as active rating

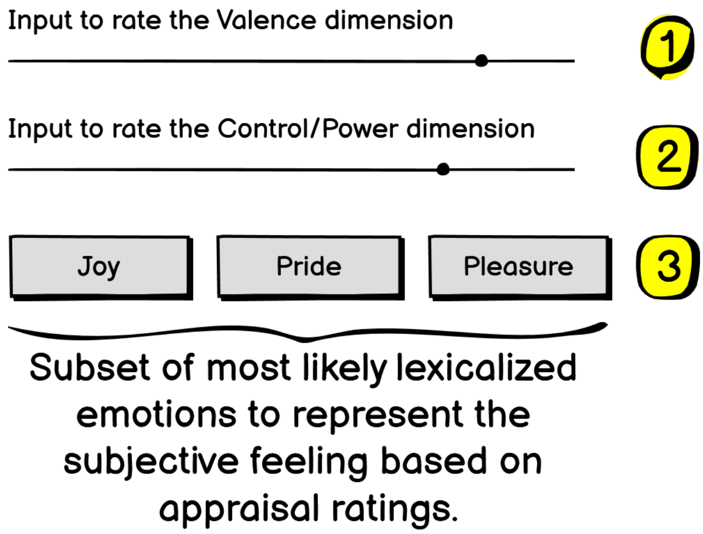


Figure 4.3: Wireframe sketching the core principle of the Dynamic Emotion Wheel in three sequential and coordinated steps.

inputs. If how a person evaluates the situation on appraisal criteria is the determinant of elicitation and differentiation of the emotional response, than it is in both the intra-personal and inter-personal best interests that the cognitive evaluation of an event is explicitly manifested through the tool. Obviously, being a self-report measure, the ratings on the appraisal dimensions are only a proxy of the *true* profile of the sequential evaluation checks. It also means that, referring to the CPM's four levels of processing (sensori-motor, schematic, association, and conceptuel), the appraisal is performed at the conceptual level, the one which requires consciousness of the underlying evaluation and effort-full calculation in the pre-frontal cortex (Scherer, 2013b), see also Section 3.4.2.

The first step consists thus in evaluating the *Valence* dimension, which spans the *Relevance* and *Implications/Consequences* appraisal groups, particularly with respect to the intrinsic pleasantness and goal conduciveness sequential evaluation checks. As stated above and throughout the theoretical Part I, *Valence* is a widely adopted dimension in many theories of emotion, and roughly defines whether a situation is pleasant or unpleasant, even though many forms of *Valence* exist (Erbas et al., 2015; Shuman et al., 2013). The CPM more specifically identifies the intrinsic pleasantness sequential evaluation check as pleasure vs. pain detector, whereas the goal conduciveness check relates more thoroughly to whether the event hinders or facilitate the attainment of the agent objectives (Scherer, 2005, 2009b).

The second steps consists in evaluating the *Control/Power* dimension, which is a proxy for the CPM's *Coping potential* appraisal group. The coping potential has been a pivotal concept from the very onset of appraisal theories (Lazarus, 1966; C. Smith & Ellsworth, 1985). It roughly refers to the extent by which the person can deal with a particular situation (Scherer, 2005, 2009b). The CPM more specifically breaks down the *Coping potential* group in three sequential evaluation checks: (1) the *control check* dictates whether the agent can be a principal actor in determining events; (2) the *power check* quantifies the effort needed to modify the course of events if the previous check gave a positive outcome; and (3) the *adjustment capacity check* relates to the agent's capacity to accept the consequences of the event, for instance if its control and power are limited (*ibid.*).

The *Valence* dimension is rated before the *Control/Power* dimension as a means to mimic the unfolding process determined by the sequential evaluation checks of the CPM. As a reminder, the CPM posits that the sequential process is obtained by appraisals ranked first in the list reaching preliminary closure before that information is conveyed to subsequent checks. In this regard, appraisals belonging to the *Relevance* and *Implication/Consequences* groups (coalescing into *Valence*) are evaluated before checks in the *Coping potential* group (representing *Control/Power*). So, even though the respondent is self-reporting a process that has already taken place, it could be useful to maintain the appropriate order as a means to enhance self-reflection on how the emotional episode was enacted.

The third step concerns the subjective feeling and consists in choosing a lexicalized emotion that best represents the whole emotional process. This is where the *dynamic* happens. Once the cognitive evaluation of the situation has been rated, the DEW builds on the probabilistic link between appraisal dimensions and the subjective feeling by suggesting a subset of discrete/lexicalized emotions that are *most likely* to occur, given the particular appraisal profile at hand. To determine which discrete emotions are most likely to occur, the DEW uses a parsimonious computational model – described in details in Chapter 5 – that considers the underlying affective space as a *reference frame.* Taking the structure of the GEW as an example, if the person rates the situation as pleasant (positive *Valence*) and thinks she is able to modify it if desired (high *Control/Power*), than the subjective feeling is most likely to be represented by one of the lexicalized/discrete emotions in the top-right quadrant of the GEW, namely *interest*, *amusement*, *pride*, *joy*, or *pleasure* (see Figure 4.2). At the same time, depending on the specific values of the *Valence* and *Control/Power* dimensions, labels adjacent to the quadrant such as *anger* (top-right quadrant) or *contentment* (bottom-right quadrant) are also more likely to occur than, say, *regret* or *guilt*, which are at the *antipodes* of the affective space. For instance, if the person evaluates her coping potential to be elevate, but is less sure about whether the situation is pleasant or unpleasant, *anger* is probably more likely to occur than *pleasure*, which is lower in *Control/Power* but more positive on *Valence*.

Finally, the mechanisms must also reckon its limitations. For instance, the appraisal is only a proxy, and also a limited proxy since it selects a partial number of sequential evaluation checks, coalescing them in two overarching dimensions. Moreover, the suggested lexicalized emotions are based on a probabilistic model, which is additionally contingent to the *adequate* theoretical appraisal of the situation (see Section 3.5.1 on the appraisal competence). In other words, the respondent interpretation of the *Valence* and *Control/Power* dimensions should be aligned with the theoretical dimensions, which determine the disposition of the lexicalized emotions in the affective space. For these reasons, but also to comply with the GEW interface and good practices in emotion self-report more generally (Mortillaro & Mehu, 2015), the third step also include a "No emotion" and an "Other. . . " field. In this way, the system gives respondents the possibility to by-pass the suggested options and provide a custom response. Figure 4.4 depicts the additional elements.

The core principle of the DEW consists thus in projecting an emotion structure into

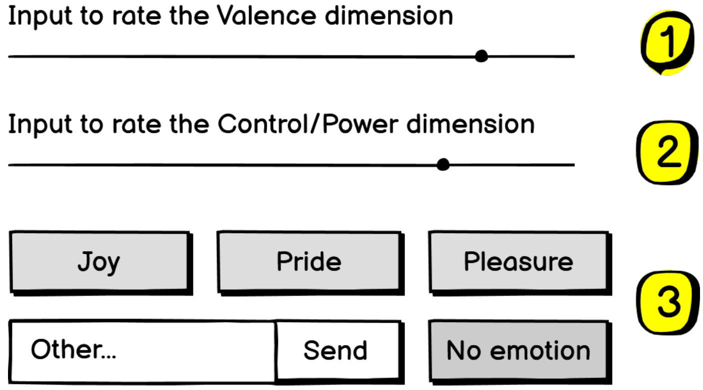


Figure 4.4: Extension of the core concept with a "no emotion" and "Other..." options.

the way the emotion is self-reported. Following appraisal theories principles, and the CPM more specifically, the emotion structure is represented by the Appraisal module and the Integration/Categorization module (see Section 3.4 and Figure 3.2 in particular). The Appraisal module is represented by the cognitive evaluation on the *Valence* and *Control/Power* dimensions, also adopted by the Geneva Emotion Wheel. The Integration/Categorization module is represented by the choice between lexicalized emotions provided by the system, a "no emotion", and a custom response. The lexicalized emotions suggested by the system are only a subset of the list pre-defined in the underlying affective space. This subset is determined by a parsimonious computational model based on the probabilistic link between the cognitive evaluation of an event and the form by which the resulting experience, the subjective feeling, can be integrated and categorized using lexicalized emotions. This is consistent with the CPM's assumption reported from Section 3.4.4 and empirically validated by a number of studies (Fontaine et al., 2007; Fontaine et al., 2013; Gentsch et al., 2017; Gillioz et al., 2016; Scherer & Fontaine, 2018; Scherer & Meuleman, 2013), and which is well resumed in this sentence:

> if one knows the results of an individual's event appraisal on major checks, one can approximately predict what kind of emotion he or she will most likely experience (or more precisely, what label the person is likely to use to refer to the experience).
> – Scherer (2009b), p. 1326

To sum up, the quest for defining the *unit* of measure to encode and convey emotional awareness that has started from broad emotion theories has now come to an end. An emotion for the DEW is represented by the value on two appraisal dimensions, combined with the subjective feeling in the form of a lexicalized emotion which can be either part of a predefined set, or provided by the user.

4.3.2 The Principle Implemented Into a Graphical User Interface

The core principle of the DEW paves the way, conceptually, for the system to prompt the *unit* of measure to encode and convey emotional awareness. As it is the case in user experience design, though, concepts must find the appropriate correspondence in the Graphical User Interface through which the user interact with the system. Since the DEW was originally conceived as a means to provide inter-individual emotional awareness, the user interface must provide both the expressing-displaying and the perceiving-monitoring functions of awareness.

The former is almost charted by the core principle. The low-fidelity wireframe of Figure 4.4 already provide a good idea of how emotion can be self-reported. An important missing part, though, concerns the specific way by which the appraisal dimensions are labelled. *Valence* and *Control/Power* are scientific terms, unusual in everyday language. Both would therefore require translation in more intuitive and comprehensible forms. As described in the attempts to validate the disposition of lexicalized emotions with direct rating for the GEW in Section 4.2.3, this task may prove to be complicated, in particular considering the requirement for the tool to occupy a limited share of the screen. This excludes long description of the dimension as the ones adopted in direct rating for the GEW (e.g., "the situation is experienced as (un)pleasant and enjoyable (disagreeable) and/or is likely to have positive and desired (negative and undesired) consequences for the person").

The perceiving-monitoring function, on the other hand, needs to be almost invented. As highlighted in Section 2.4, scant attention has been given so far to how emotion can be graphically represented in a computer-mediated (learning) environment (Derick et al., 2017; Ez-zaouia et al., 2020; Leony et al., 2013). The building blocks for decoding emotional awareness are therefore the information that has been encoded: the value on appraisal dimensions and the subjective feeling.

The attempt to integrate the different aspects in a comprehensive user interface is depicted in Figure 4.5. The interface occupies around a fifth of the width and the whole height of a *modern* screen (depending on the resolution). It is divided in two main areas: the upper part is devoted to the expressing-displaying function, whereas the bottom part to the perceiving-monitoring function of awareness tools.

The expressing-displaying function implements the core principle as described by the wireframes, using a uniform orange color for most of the elements of the interface that the user can manipulate. The two appraisal dimensions are represented by two horizontal sliders, each one with a continuum ranging from a negative *not at all* to a positive *yes, absolutely* poles. The neutral point is highlighted with a cue in the middle of the slider range. The *Valence* dimension is prompted by the label *VALENCE - Is the situation pleasant?* whereas the *Control/Power* with the label *POWER - Is the situation under your control?*. The choice of wording is adapted from the complementary material available with the Geneva Emotion Wheel. The name of the dimension in uppercase has been provided as a cue for the appraisal dimensions in the lower part of the screen, described below.

The suggested subjective feelings are represented by three orange buttons, labelled with the lexicalized emotions that are most likely to occur given the values of the two appraisal-sliders. Below the three buttons, a text-input allows user to enter a custom response, which must be confirmed with a *Send* button. When the field is activated, a drop-down menu shows all the available lexicalized emotion. Beside the custom response field, a *No emotion* completes the options to encode an emotional episode into the tool. The overall interaction

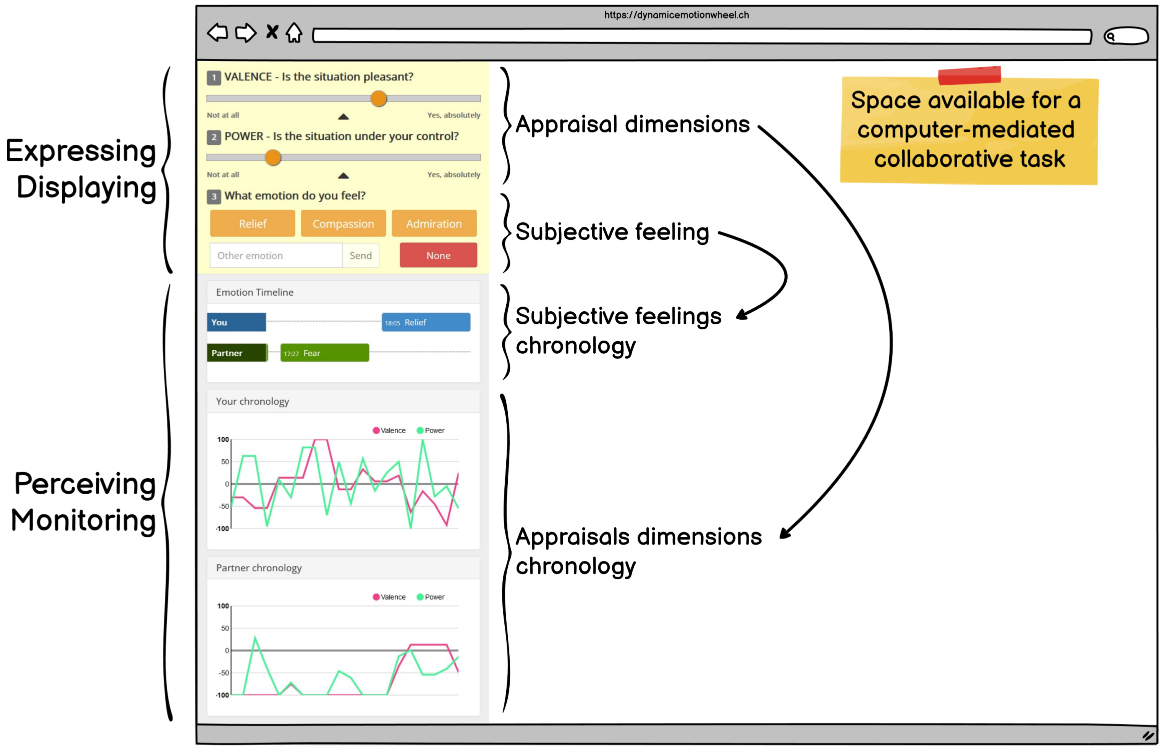


Figure 4.5: The overall interface of the prototype, with visual cues about the role of each element. The size of the tool is not representative in proportion to the browser's window. The actual size is closer to one-fifth of the width.

workflow is depicted with three screen captures disposed side-by-side in Figure 4.6.

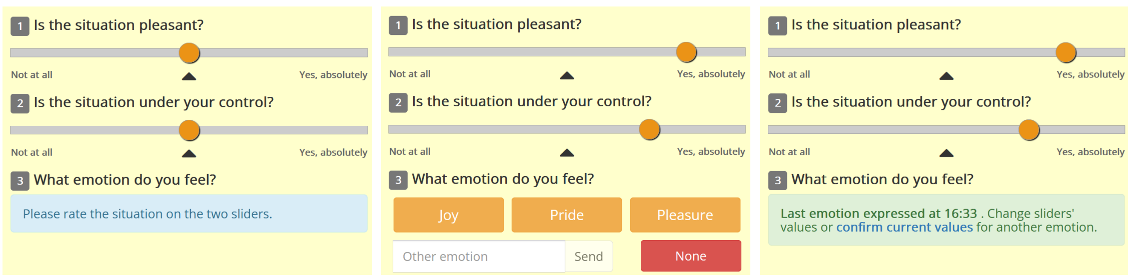


Figure 4.6: The sequence to express-display an emotion through the Dynamic Emotion Wheel.

The user is first prompted to rate the appraisal sliders. This message appears only the very first time the DEW interface appears and covers the space where the subjective feelings will appear. In this way, the dynamic nature of the link between the sliders and the buttons is perceived from the very onset of the interaction. The two sliders are also numbered to enforce the rating order corresponding to the sequential evaluation checks. As soon as the user moves one knobs of the appraisal sliders, the options to express the subjective feelings are computed istantly. The number 3 besides the appearing zone completes the sequence of steps necessary to express the emotion. The user can at this moment observe the computed subset of discrete emotions and assess whether any of them is suitable to represent the actual subjective feeling. In that case, the corresponding button can be clicked. Otherwise, the "Other.." field or the "No emotion" buttons can be used instead. Once one of the options to express the subjective feeling is triggered, the system confirms the emotion has been encoded, showing also the time of the action. The confirmation message also informs the user that, in order to express another emotion, the sliders must be moved again. A link can be clicked if the cognitive evaluation hasn't changed. This feature is intended especially in the case of blended or mixed emotions (Scherer, 2005; Scherer & Meuleman, 2013), which can be expressed in rapid succession by modifying only the subjective feeling, leaving the appraisal dimensions at the same values.

The perceiving-monitoring part of the screen mirrors the emotion structure encoded by the system. The upper part is devoted to the subjective feeling, whereas the bottom part to the appraisal dimensions.

For the subjective feelings, an emotion timeline shows the chronology of the lexicalized emotion for the person herself (top) and the partner (bottom). The lexicalized emotions are added on the right of the line, pushing the older ones back. The two lines are synchronized, so that the last visible emotions are those more recent in time, regardless of whom has expressed them. In Figure 4.5, for instance, the last subjective feeling is *Relief* and was encoded by the person using the tool, whereas the second-to-last is *Fear*, expressed by the partner. A scrolling bar allows participants to scale back and forth if desired. The aim of the emotion timeline is to provide both an up-to-date and a chronological comparison of the emotional experience between the person and the partner.

Below the emotion timeline, two line charts report the chronology of the appraisal dimensions: the top one for the person herself, the bottom one for the partner. Each dimension has its own line on each chart. The lines also shrink and adapt when a new emotion is inserted into the system, even though each chart shrinks independently from the other. In Figure 4.5, we can for instance see that the line chart for the person using the tool is more nuanced compared to the partner, whose appraisals tend to jump from one extreme of the continuum to the other, and also overlap considerably. The aim of the charts is to pass into perception the evolution of the cognitive evaluation over time (Hegarty, 2011).

To sum up, the DEW thus maintains most of the theoretically and empirically validated characteristics of the Geneva Emotion Wheel, but adapting and extending some functions to better fit the perspective of an awareness tool rather than a self-report tool alone. An important missing feature compared to the GEW is the possibility to rate the intensity of the emotional episode. The values of the appraisal dimensions are in fact not to be confused with the intensity of the emotion.

Furthermore, the interface allows to meet most of the requirements set forth with the EATMINT project's members, listed in Section 4.1.2. First, the tool clearly maintains its self-report perspective. Second, it implements an emotion structure, based on the Component Process Model theoretical framework. Third, it reduces the number of concurrent choices on the screen, leveraging on the dynamic mechanisms of sub-setting the underlying affective space. Fourth, the space gained from a reduced expressing-displaying part is used to provide graphical representation of emotions, which are meant to increase the perception of the evolution of the emotional experience over time, as well as foster comparison between the person and the partner. Finally, by abstracting the computational model and harnessing the construction by modules of the graphical representation, it is also possible to adapt the inner-functioning and/or the appearance of the tool. In this regard, the next section shows a first adaptation that was made by creating an underlying affective space starting from the discrete emotion used in the EATMINT tool.

4.3.3 The EATMINT Circumplex

A first theoretical, rather than technical, distinction from the Geneva Emotion Wheel consisted in adapting the underlying affective space to lexicalized emotions more likely to occur in computer-mediated collaborative tasks. The 20 lexicalized emotions of the GEW have been in fact chosen for their relative frequency in everyday experience at large (Scherer, 2005; Scherer, Shuman, et al., 2013; Shuman & Scherer, 2014). The EATMINT team had already provided a list of discrete emotions listed in Table 4.1 of Section 4.1.1. It was thus possible to exploit this list, provided that each lexicalized emotion could be placed on an underlying circumplex.

As suggested by the authors having worked on the GEW, one valid way to perform this task would have been the use of the GRID instrument (Fontaine et al., 2013; Scherer, Shuman, et al., 2013). Nevertheless, setting up a study asking participants to rate the discrete emotions seemed too risky without a prior test of the prototype. As a consequence, the discrete emotions used by the EATMINT team were placed by searching correspondences in existing bi-dimensional affective spaces. In particular, the four circumplex-like affective spaces were identified: versions 2.0 and 3.0 of the Geneva Emotion Wheel (Scherer, Shuman, et al., 2013), the circumplex model of affect James A. Russell (1980), and the alternative dimensional structures of the semantic space for emotions derived by Scherer (Scherer, 2005), which also integrates the circumplex from the pan-cultural study by Russel (James A. Russell, 1983). Whereas the former two affective spaces use the *Valence* x *Control/Power* dimensions, the latter use the *Valence* x *Arousal* dimensions. In this regard, though, Scherer (2005, p. 722) notes that "a 45° rotation of the axes [of Russell's circumplex] corresponds rather nicely to an explanation of the distribution of the terms in a two-dimensional space formed by goal conduciveness and coping potential".

Using these sources and conceding a certain degree of approximation due to the translation, it was possible to find a match for 16 out of the 20 EATMINT discrete emotions. Table 4.2 lists the original EATMINT label, the label and the affective space of the source, and the value on the *Valence* and *Control/Power* dimensions.

Two other discrete emotions had a partial matching in the sources. *Anxious* could be related to *Worry* (GEW 2.0) or *Anxious* (Russel, 1980). In one case, though, the *Valence* was positive and in the other negative, whereas the *Control/Power* was negative in both sources. *Relaxed* was matched with *Relaxed* or *Calm* in Scherer (2005), in which case the *Valence* was positive, but the *Control/Power* was positive for one and negative for the other. It was decided to place the *Anxious* in the quadrant with both negative values, whereas *Relaxed* was placed in the positive *Valence* and negative *Control/Power* quadrant. With this configuration, the 18 discrete emotions placed so far formed a nicely balanced

Table 4.2: Matching in existing affective spaces to create the EATMINT circumplex

EATMINT	Source	Affective space	Val.	C./P.
Amused	Amusement	GEW 3.0	+	+
Annoyed	Annoyed	Scherer, 2005	-	+
Attentive	Attentive	Scherer, 2005	+	+
Bored	Bored	Scherer, 2005 or Russel, 1980	-	-
Confident	Self-confident	Scherer, 2005	+	+
Delighted	Joy or Elation	GEW 2.0 or 3.0	+	+
Disappointed	Disappointed	GEW 3.0	-	-
Empathetic	Compassion or Empathetic	GEW 3.0 or Scherer, 2005	+	-
Envious	Envy	GEW 2.0	-	+
Frustrated	Frustrated	Scherer, 2005	-	-
Interested	Interest	GEW 3.0	+	+
Irritated	Irritation	GEW 2.0	-	+
Relieved	Relief	GEW 3.0	+	-
Satisfied	Contentment	GEW 3.0	+	-
Stressed	Stressed	Russel, 1980	-	+
Surprised	Surprise	GEW 2.0	+	-

circumplex, with 5 emotions in the top-right and bottom-right quadrants, and 4 emotions for the quadrants on the left side.

The two lasting discrete emotions did not have even a partial matching. So it was decided using an heuristic approach to place *Confus*ed in the quadrant with both negative dimensions. The very last emotion, *Grateful*, seemed more adequate in the positive *Valence* and negative *Control/Power* quadrant, but this quadrant tallied already 5 emotions. The only quadrant with a missing discrete emotion was the negative *Valence*, positive *Control/Power*. It was thus decided to leave *Grateful* out of the circumplex, and replace it with *Dégoûté*, translated by *Disgusted*. The final disposition is depicted in Figure 4.7.

The resulting EATMINT circumplex has a double interest. On a practical standpoint, it is an affective space that is meant to provide users with lexicalized emotions tailored to computer-mediated collaboration, which may be extended to Computer-Supported Collaborative Learning (see Section 1.3.4). On a more abstract level, it is a first step in the customization of the Dynamic Emotion Wheel, which can be driven by theoretical and/or empirical foundations. Before heading this subject more thoroughly, the results of a usability test are illustrated to provide a preliminary assessment of the prototype.

4.3.4 Usability Test of the Dynamic Emotion Wheel

The prototype was ultimately tested with 16 participants, recruited in the hall of Geneva University, who took part to the test without any remuneration. Participants were university students at the undergraduate and graduate levels. The test adopted an experimental setting in which participants were lead to believe they were collaborating in a problem

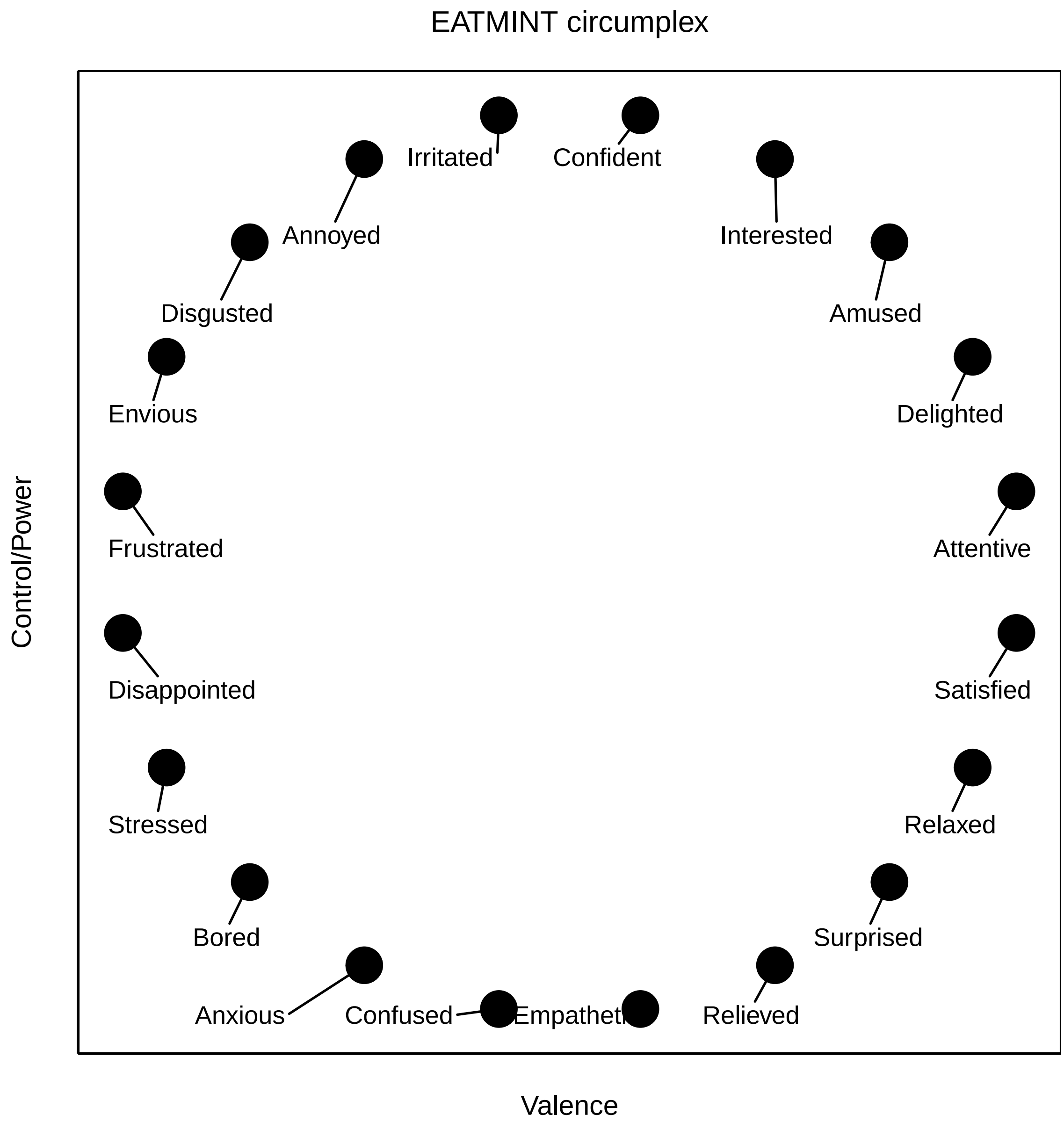


Figure 4.7: Disposition of 19 out of 20 discrete emotion used in the EATMINT project, with *Grateful* replaced by *Disgusted* to balance 5 items in every quadrant. The result is the EATMINT circumplex.

solving task, whereas in fact, they were interacting with the *playback* of a confederate that had undergone the same task, but with another confederate. The same controlled environment will be reused for the first empirical contribution, so the details of the test can be retrieved from the Methods section of Chapter 7. The results of the test will also be integrated in the third empirical contribution of Chapter 9. A thorough depiction of the test would therefore duplicate the information. A brief summary of the main findings is therefore outlined, whereas the next subsection illustrates some measures that will be of interest for the empirical contributions of the thesis.

The test mainly collected three types of information: the concrete use of the DEW in expressing emotions during the task, eye-tracking measures that recorded how the participant interacted with the whole interface (Poole & Ball, 2005), and measure about the perception of the tool's usability (Brooke, 1996). Overall, the test provided encouraging results, but also highlighted some shortcomings.

With respect to the expressive-displaying function, participants were asked to express at least 3 emotions for each of the 4 problems to solve. They eventually expressed around 17 emotions during a task of 20 minutes, so more than the demanded number. They found the mechanisms relatively easy to use. The data, though, also showed that the two appraisal dimensions tended to be rated symmetrically, a problem that will be thoroughly addressed in Chapter 9.

The overall interface of the test was divided in three areas of interest. First, the interface was horizontally split in two. On the left-hand side of the screen, the EAT disposed of two areas of interest corresponding to the expressing-displaying and perceiving-monitoring parts respectively. The right-hand side covered the part of the interface dedicated to the simulated problem-solving task. The overall interface can be inferred from the heatmap illustrated in Figure 4.8, which also depicts which elements have been perused longer by participants. Two kind of eye-tracking measures were retained for the analysis: the number of visits inside each area of interest, and the total number of seconds that participants spent perusing each area. The first is considered a measure of information seeking, whereas the second of information processing (Poole & Ball, 2005). Data showed that participants spent more of the time in producing emotional awareness rather than perusing it. When the emotional information was attended by, the emotion timeline with the subjective feelings was preferred to the line charts depicting the evolution of the cognitive evaluation.

Finally, at the end of the task, participants filled a survey where they were asked about their perception of the EAT. The survey comprised also the System Usability Scale (Brooke, 1996), a widely adopted scale to measure usability, which has incidentally been administered also by Feidakis et al. (2014) to assess the *emot-control* tool already mentioned (see the related works in Section 2.4). The overall rating on the scale was of almost 75 points out of 100, which can be considered an acceptable usability score (Bangor, Kortum, & Miller, 2009; Lewis & Sauro, 2018). A thorough assessment of the usability score will be discussed in Chapter 9.

To sum up, the usability test provided interesting feedback for the DEW. It also highlighted how the use and perception of the tool varied from participant to participant. In the open-ended discussion at the end of the test, participants confirmed they found the

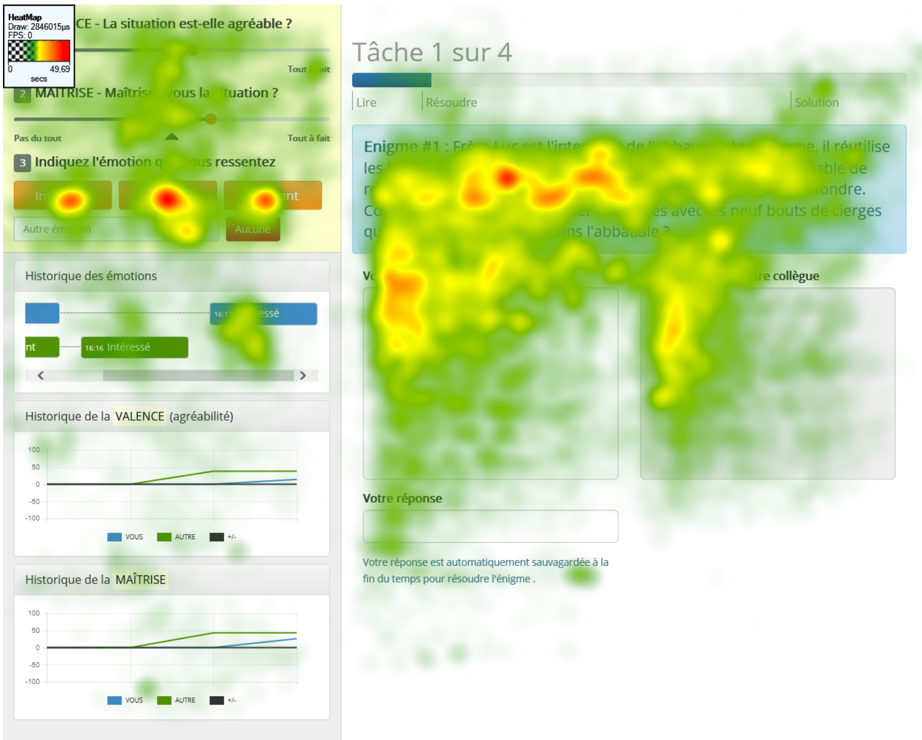


Figure 4.8: Heatmap representing the time the gaze has lasted on each part of the interface. Red zones represent long-gazed elements.

overall concept intriguing, even if they have never thought of expressing emotions during a computer-mediated collaborative task. In the meantime, most of them also manifested some skepticism on whether they would adopt the tool in more formal settings.

4.3.5 Retained Measures for Further Empirical Contributions

In this subsection, a selection of measures observed in the usability test are grouped into a reference table. Due to technical problems, not all the measure are retained for the $N = 16$ participants who took part to the test. The number of participants each measure is based upon is therefore reported alongside the mean and standard deviation. The measure have been recomputed since the Master thesis, and are therefore to be considered more reliable in case of discrepancy. These measures will be used in analysis and comparisons in the empirical part of the thesis.

Table 4.3: Reference measures from the usability test in Fritz (2015)

Description	N	M	SD
Visits duration in the Task Area (seconds)	12	650.61	116.53
Visits duration in the EAT Area (seconds)	12	231.55	70.86
Number of emotion expressed by participant	14	17.14	5.05
Visits duration in the Perceiving-Monitoring Area (seconds)	12	57.32	21.67
Visits count in the Perceiving-Monitoring Area	12	82.25	25.51
Overall score on the System Usability Scale	14	74.88	14.81

4.4 Extending the Prototype Into a Proof of Concept

The chapter has so far covered how the Dynamic Emotion Wheel came to be as an attempt to respond to specific requirements within the EATMINT project scope. As a consequence, it was highly oriented toward the inter-personal function of an awareness tool, especially in a dyadic interaction. As illustrated in Chapter 2, though, emotional awareness can be thought of also at the intra-personal level. In addition, the inter-individual level does not have to be constrained to dyads neither. Furthermore, the use of the EATMINT circumplex highlights how the DEW can be modified according to theoretical and empirical criteria. In the reminder of the chapter, so, a series of features are set forth to transform the DEW from a *single-case* prototype to a multi-purpose proof of concept.

As a reminder from the introduction, the term *proof fo concept* refers here to a functioning version of a product that has focused primarily on conceptual features rather than technical reliability. It is also worth reminding from the introduction that in this Part II, the EAT is presented in its latest version from the iterative design process. This means that some features have been added following the results of the empirical contributions in Part III, and also that those contributions did not dispose of all the same features presented in this part.

The proof of concept mainly consists in two integrated elements. First, the parsimonious computational model based on the probabilistic link between appraisal dimensions and subjective feelings. The next chapter will describe this model at the core of the DEW, in particular the fact that it does not have to be restrained to affective spaces characterized by two dimensions. Second, a toolbox built around the core computational model, allowing researchers and practitioners to (1) set up their own instance of the tool – by determining autonomously important affective and/or pedagogical choices – and (2) share configuration and data from an open-science perspective. Chapter 6 will describe the main features of the toolbox and how they can be adopted to various aims, including facilitating the comparison between researches on the subject of emotional awareness.

4.5 Summary

This chapter provided an overview of the Dynamic Emotion Wheel, a prototype built in the specific context of the EATMINT project, but that has, as a result of its design, an inner functioning that can be easily abstracted to different theoretical framework, research settings, and pedagogical uses. The tool is inspired by the Geneva Emotion Wheel, a self-report instrument deeply rooted in the Component Process Model. The core principle of the DEW consists in twisting the function of the appraisal dimensions: whereas in the GEW they are adopted as axes on which lexicalized emotion are disposed, in the DEW the user actively rate the *Valence* and the *Control/Power* dimensions in order to obtain a subset of most likely subjective feeling to occur, given the specific evaluation at hand. The emotion structure consisting in the measure of the two appraisal sliders and the subjective feeling is then used for conveying emotional awareness through a timeline of subjective feelings and line charts for the appraisal dimensions. A test of the interface conducted in a simulated computer-mediated collaborative environment yielded mixed, but all things considering encouraging results. The two next chapters will describe how the DEW has been theoretically and technically extended to better meet practitioners and researchers' needs.

Chapter 5

A Parsimonious Computational Model Linking Appraisal and Subjective Feeling

This chapter illustrates the parsimonious computational model that is at the core of the Dynamic Emotion Wheel. The model is based upon the probabilistic link, stated by appraisal theories of emotions and the Component Process Model more specifically, between the cognitive evaluation of the situation and the resulting holistic experience, the subjective feeling, that can be lexicalized in natural language words or idioms. The combination of the computational model and the interface of the DEW illustrated in the previous chapter makes it possible to abstract from two-dimensional spaces and take advantage of underlying structure for which the number of dimensions can be defined according to researchers and practitioners' choice. In this regard, the chapter starts by briefly outlining the interest for affective spaces with a different number of underlying dimensions. Second, it illustrates the logic behind the computational model inside the DEW. The following three sections, then, describe how the same computational logic can be applied to one-dimensional, two-dimensional, and multi-dimensional affective spaces (3 and beyond). The chapter ends with an assessment of the limitations of the computational model.

5.1 Interest For N-Dimensional Affective Spaces

As the use of the EATMINT circumplex depicted in Section 4.3.3 has shown, it is possible to adapt the underlying *reference frame* that link the appraisal dimensions with the subjective feelings. Still, the example remained in the limits of a bi-dimensional affective space. The use of this kind of affective space, though, was driven mainly by the familiarity and widespread use of the graphical representation in the literature as well as in other self-report tools, in particular the Geneva Emotion Wheel. Beyond the theoretical foundation of bi-dimensional affective spaces, the representation also happens to be the highest number of dimensions that can be easily represented graphically on a screen or sheet of paper. Nevertheless, once the graphical user interface of the Dynamic Emotion Wheel breaks the *overlapping* between the affective space and the self-report tool, than a circumplex becomes one *reference frame* among others. In other words, the circumplex determines only the underlying theoretical and/or empirically validated affective space, but does not force the interface to change dramatically. In fact, the appraisal dimensions in the DEW are *flattened* into the rating of the sliders. Concretely, more appraisal dimensions would simply translate in more sliders through which respondents rate the situation. Furthermore, the computational model can subset only a limited number of subjective feelings, which does not require thus that all of them are represented at the same time. This fact has several consequences both from an *affective science* and an *emotional awareness* perspectives, which are considered hereafter.

5.1.1 Affective Science's Interest in *N*-Dimensional Spaces

As it has been shown in Part I, researchers adopt very different conceptualization of emotion across and within the same family of emotion theories (see Section 3.2). Theories assuming that a dimensional mechanism is implicated in the structure of emotion assume different positions on the number of dimensions, ranging from a uni-dimensional structure (*positive* vs. *negative* emotions) to the multiple sequential evaluation checks of the Component Process Model listed in Section 3.4.2.

In this regard, empirical evidence gathered in the last decade has also highlighted how two dimensions are not enough to account for the variety of emotional experiences, even though some of them have more *weight* than others in discriminating between discrete emotions (Fontaine et al., 2007; Fontaine et al., 2013; Gentsch et al., 2017; Gillioz et al., 2016). In an attempt to provide a working definition of emotion based on seven defining features, Mulligan and Scherer (2012, p. 346) propose that "x is an emotion *only if* [. . .] x is triggered by at least one appraisal, and x is guided by at least one appraisal". The working definition clearly states that one appraisal is necessary, but does not provide any fixed number or upper-limit.

Even when the same number of dimensions are postulated, there is no agreement on what these dimensions should be, across and within families of emotion theories. Taking two-dimensional affective spaces as an example, we can identify at least (1) the ubiquitous *Valence* x *Arousal* dimensions (James A. Russell, 1980; James A. Russell, 1983; Stanley & Meyer, 2009) used for instance by the *emot-control* (Feidakis et al., 2014; Feidakis et al., 2013) and Mood Meter application (Brackett et al., 2019; Hoffmann et al., 2020; Nathanson et al., 2016); (2) the *Valence* x *Control/Power* dimensions of the Geneva Emotion Wheel (Scherer, 2005; Scherer, Shuman, et al., 2013; Shuman & Scherer, 2014); or (3) Pekrun's *Control* x *Value* dimensions in achievement emotions (Pekrun, 2006; Pekrun & Perry, 2014).

At the same time, the number of lexicalized emotions representing the subjective feelings does not have to be bound to 20 or any other fixed number. Once again, it should be up to the researcher or practitioner to decide whether emotion differentiation between a small or large number of options is an important part of the process (Erbas et al., 2015; Moors et al., 2013). For instance the EMORE-L tool by Gaëlle Molinari et al. (2016) comprises 8 discrete emotions, whereas the digital version of the Mood Meter application linked to the RULER approach (Brackett et al., 2019; Hoffmann et al., 2020) allows respondents to choose between 100 lexicalized emotions.

To sum-up, the DEW inner functioning not only allows to break through a forced bi-dimensional affective space, but it also permits different dimensional affective spaces to coexists within more or less the same interface. This may be interesting to assess and compare the contribution of a lower vs. a higher-order of dimensions in active rating.

5.1.2 Emotional Awareness's Interest in *N*-Dimensional Spaces

The DEW adopted two dimensions for usability constraints: it was considered important that expressing an emotion would not take too much time and effort. As described in Section 1.3.2, though, there is a growing consensus in considering that the time and effort in producing and consuming awareness should not be evaluated in absolute terms, but in a relative trade-off, which accounts for the instrumentality of the information provided in helping learners to improve their learning processes and outcomes (Buder, 2011; Dillenbourg et al., 2016; Janssen & Bodemer, 2013). The number of dimensions should therefore be a pedagogical decision, not imposed *a priori* by the interface of the tool. For instance, in the EMORE-L tool (Gaëlle Molinari et al., 2016), learners use three dimensions to rate their emotional experience: the subjective control over the situation, the subjective value attributed to the event, and the perceived level of activation. As outlined in the abstract model of emotional awareness in Section 2.5, an EAT can help learners in extrapolating meaning-making by guiding their introspection and self-reflection. Appraisal criteria – or dimensions more broadly – may have a privileged function in this sense, and it would be limiting to have to decide between a fixed number of them for visuo-spatial constraints.

As for the *affective science* perspective illustrated above, the same logic applies to the pre-determined lexicalized emotions that are presented to learners. Depending on the accuracy and level of granularity desired in representing the subjective feeling, the underlying affective space can provide any number of options. From a computational perspective, these options would be all the more distinct if they are *scattered* in multiple dimensions. If the use of an EAT is, for instance, directed towards improving the ability of learners in discerning between similar, but not identical, emotional experiences, having a tool that provides adjacent lexicalized emotions as a function of multiple evaluation can be a means to orient learners into fine-grained differences. The subset mechanism of the DEW provides in this regard the possibility to maintain a sheer number of options, without having them all together on the interface.

A higher number of dimensions in self-reporting emotions also provide more accurate information for the graphical representation in the perceiving-monitoring function of awareness. From a reverse-engineering perspective, it is easier to answer the questions "what does my colleague mean when she says she feels x?", when x can be quantified on different appraisal dimensions. For instance, with *Valence* alone, the learner can only infer if x is roughly pleasant or unpleasant. With *Valence* and *Control/Power*, the learner can infer (un)pleasantness and the extent by which the colleague thinks she can act upon the situation. Finally, with *Valence*, *Control/Power* and, say, *Agency*, the learner can infer all of the above, but also obtaining cues about whether the causes of x are to be looked into the actions of the colleague, or if the learner herself is somehow responsible.

Once again, the decision of striving for immediacy or accuracy in self-reporting emotion must be assessed in terms of pedagogical implications, rather than available user interfaces. A computational model that allows to make this decision by minimizing the impact on the technical requirements is therefore also pertinent to an *emotional awareness*'s interest. The next section thus describe the logic of model.

5.2 The Foundations of the Computational Model Based on a *k*-Nearest Neighbors Approach

This section first proposes a brief and general outline of computational models of emotion, used to situate the model at hand. Then, the model is described in terms of input, treatment, and output. Finally, the section lays out the requirements for the model to be adopted. Both description and requirements are indicated at a conceptual, abstract level. The concrete implementation in a programming language is available in the Code and Data Section at the end of the Introduction.

5.2.1 Computational Models of Emotion

Computation models of emotion are a wide domain of inter-disciplinary research, spanning fundamental emotion theories and highly applied applications with computational agents endowed with emotions (Calvo & D'Mello, 2010; Guest & Martin, 2021; Marsella et al., 2010b; Mortillaro & Mehu, 2015), see also Section 1.2 about affect-aware systems. Compared to emotion theories more generally, which may sketch some inner functioning of emotion in broad terms while waiting for further evidence, computational models must instantiate the mechanisms in concrete terms, which are interpreted or compiled by a machine (*ibid.*). A positive turn of this necessity is that computational model can be presented – and thus scrutinized – in the smallest details.

Marsella et al. (2010b) identify three main areas to which computational models of emotion may contribute. First, they can have an impact on fundamental theories of emotion, for instance from the psychological perspective. Second, they can empower the field of artificial intelligence and robotics. And third, they can contribute to human-computer interaction both at the intra-personal and inter-personal levels. The computational model set forth in this chapter is particularly driven from this last perspective, but may also contribute – to a lesser extent though – to the first. Furthermore, Marsella et al. (2010b) also suggest that computational models, especially from a componential driven perspective, should make clear which components are included in the model. As emerged in the previous chapters of the thesis, the computational model adopted by the DEW is focused on the appraisal and categorization/integration modules.

Finally, congruently with Cernea & Kerren (2015) list of emotion-related technologies, the computational model is meant to be parsimonious by requiring minimal hardware and also limited computational power.

5.2.2 Model's Description

The aim of the model is to first sort, and later subset, a number of lexicalized emotions, which are provided as representative of the *true* emotional episode. To obtain this result, the computational model is based upon the following parameters, which are presented hereafter in a notation adopted especially throughout the chapter, and occasionally in the reminder of the manuscript:

- E stands for *Appraisal* ***E****valuation*, that is, the hypothetical value that a cognitive evaluation can have on a given set of appraisal criteria $\{E_1, E_2, \ldots, E_i\}$. In concrete terms, E represents the rating of a person asked to evaluate one or more appraisal dimensions. If there was only one appraisal, *Novelty* for instance, than E assumes a value determined by the extent to which the person evaluates the event as sudden, familiar or unpredictable on E_1. In terms of the DEW interface, E refers thus to the slider(s), which prompt respondents to actively rate the situation against an appraisal/dimensional criterion. There in only one E at a given moment t, but E can be repeated at $t_1, t_2, t_3, ...$, each time assuming either the previous value or a different value.
- P stands for *Subjective Feeling* ***P****redicted Appraisal*, that is, the predicted appraisal profile associated with a given subjective feeling, with subscripts $\{P_1, P_2, \ldots, P_j\}$ representing a set of j predicted appraisal criteria. The *coordinates* of P are identified with a common label, the lexicalized/discrete emotion, which is meant to be representative of the subjective feeling. There may be any number of labelled P^* belonging to an affective space A, that is $P^a, P^b, P^c, ... \in A$. In concrete terms, each P represents the appraisal profile that is supposed to elicit the given subjective feeling. For instance, the CPM predicts that a subjective feeling represented by the word *happiness* may be related to a higher evaluation on *Novelty* appraisal groups compared to a subjective feeling represented by the word *sadness* (Scherer, 2010b). In term of the DEW functioning, P is determined by the underlying affective space A used as a *reference frame*. Contrary to E which may varies according to t, P holds constant for any given A as long as A is adopted as the *reference frame*.
- Finally, $\Delta(E, P)$ represents the distance between the Appraisal Evaluation and the Subjective Feeling Predicted Appraisal, with the *k-nearest neighbors* being the k lexicalized emotions with the lowest $\Delta(E, P)$. In other words, if $k = 5$, then, among all the options in the affective space, the 5 whose appraisal profile P minimize the distance from E represent the subset of most likely subjective feelings to occur, given the *mysterious* point E provided by the person's cognitive evaluation. $\Delta(E, P)$ is computed at any variation of E at time t, but holds constant if $E_{t1} = E_{t2}$.

The computational model therefore consists in a function that takes the three parameters E, A, and k as Input, and provides a subset of A, consisting in a set of ordered $\overrightarrow{P}$, where the order is determined by $\Delta(E, P)$, as Output. In a more succinct form:

$$f(E, A, k) \Rightarrow \mid \overrightarrow{P} \mid = k : P \in A \text{ and } \Delta(E, P) \text{ of } P^x \leq \Delta(E, P) \text{ of } P^{x+1}$$

5.2.3 Model's Requirements

The computational model has minimal specific requirements. First, it is mandatory that the cardinality of E is inferior or equivalent to the cardinality of P. That is, if respondents are asked to evaluate the situation on three appraisals/dimensions $\{E_1, E_2, E_3\}$, then the underlying affective space must be three-dimensional or higher. It must in fact provide at least coordinates for $\{P_1, P_2, P_3\}$, with the additional dimensions $P_4 \dots P_j$ not adopted as *reference frames*.

Second, it is necessary that the domain of E – the possible values of the rating – matches the domain of P. If the evaluation is rated in a domain ranging from $[V_{min}:V_{max}]$, the lexicalized emotions in the affective space must be disposed somewhere between V_{min} and V_{max}. Any values in the range can be potentially acceptable by the model, including the very same values for all options. In which case, though, the subsetting algorithm would be useless, since every lexicalized emotion would be always at the very same distance from E. But as long as at least two options in the affective space do not have exactly the same P, than the computation model guarantees to sort the suggested subjective feelings according to the minimal value of $\Delta(E, P)$.

The following sections thus illustrate how this distance $\Delta(E, P)$ is calculated by the model depending on the number of appraisal criteria/dimensions composing the underlying affective space: one-dimensional, two-dimensional, or multi-dimensional. The presence of a specific section on two-dimensional affective spaces – which are in fact the first multi-dimensional condition – is due to the possibility of interpreting the affective space either as a circumplex or as a Cartesian plane, which require a slightly different computation. The specific interest of each n-dimensional space are briefly discussed with respect to the existing literature.

5.3 One-Dimensional Affective Spaces

In a one-dimensional affective space, where there is only a single appraisal criterion, the nearest neighbors are determined by the absolute distance between E_1 and P_1:

$$\Delta(E, P) = |E_1 - P_1|$$

One-dimensional affective space are not usually defined as such in the literature, even though the possibility to conceptualize emotion on a single dimension is not infrequent, especially when this dimension represents an omnibus *Valence* (Erbas et al., 2015; Shuman et al., 2013). The possibility to use a one-dimensional affective space is also warranted by the working definition proposed by Mulligan & Scherer (2012) – *i.e.*, *at least one appraisal* – as well as by researchers and practitioners' habit to recur to a dichotomous categorization (Harley et al., 2017; Pekrun et al., 2002). Even though the actual rating may be implicit rather than explicit, for instance in grouping forced choices in a *positive* or *negative* columns as in Gaëlle Molinari, Chanel, et al. (2013), this is still a uni-dimensional space in disguise.

The possibility to restrain the affective space to a single dimension may also be useful in situations where researchers or practitioners are interested specifically in the assessment of only one criterion, avoiding thus the potential mediation or moderation effects of additional criteria. For instance, if one is interested in the *Normative significance*, more specifically with respect to *External standards* – whether the event violated laws or socially accepted norms (Scherer, 2005) – the presence of another appraisal criteria such as *Valence*, may influence that evaluation. For example, negatively valenced emotion may be considered less socially acceptable (Van Kleef, 2018). The presence of an explicit reminder of the valence could therefore mediate or moderate the rating of the external standard of normative significance.

5.4 Two-Dimensional Affective Spaces

Two-dimensional – or bi-dimensional – affective spaces are ubiquitous in research, especially in the form of the dimensional *Valence x Arousal* circumplex (James A. Russell, 1980), but also in appraisal-driven perspectives, as in Pekrun's Control-Value theory of achievement emotions (Pekrun, 2006; Pekrun & Perry, 2014) or the V*alence x Control/Power* structure of the Geneva Emotion Wheel (Scherer, 2005; Scherer, Shuman, et al., 2013). As illustrated in the first part of the thesis, several scholars have theoretically adopted two-dimensional affective spaces in their research, but also as graphical representations for self-report/awareness tools (*e.g.*, Feidakis et al., 2011). Consequently, it is important that an overarching computational model can manage to incorporate different kinds of two-dimensional affective spaces.

In this regard, an affective space based on two dimensions can be interpreted as the sum of a pair of one-dimensional continuous spaces, in which case each subjective feeling is determined by a precise coordinate in a Cartesian plane. Alternatively, it may also be interpreted, less stringently, as a circumplex where each subjective feeling is positioned at an equal distance from the origin, as in the GEW. The alternative interpretations of a two-dimensional affective space, depicted in Figure 5.1, require different computational models from a *k-nearest neighbors* perspective. This section, thus, first illustrates the computational model for a Cartesian plane, and then for a circumplex. Last, it compares the two computational models in a simulation using the same affective space depicted either as a Cartesian plane or a circumplex.

5.4.1 The Affective Space as a Cartesian Plane

When a second appraisal criteria is added, the resulting affective space can be conceptualized as a Cartesian plane, in which P is represented by the coordinates $x = P_1$, and $y = P_2$. At the same time, the appraisal evaluation is also determined by coordinates $x = E_1$ and $y = E_2$. Consequently, the *k-nearest neighbors* may be determined using the Euclidean (or Pythagorean) distance, with the nearest lexicalized emotions having the smallest hypotenuse of a triangle with the length of the sides determined by the difference between E_1 and P_1, and E_2 and P_2:

$$\Delta(E, P) = \sqrt{(E_1 - P_1)^2 + (E_2 - P_2)^2}$$

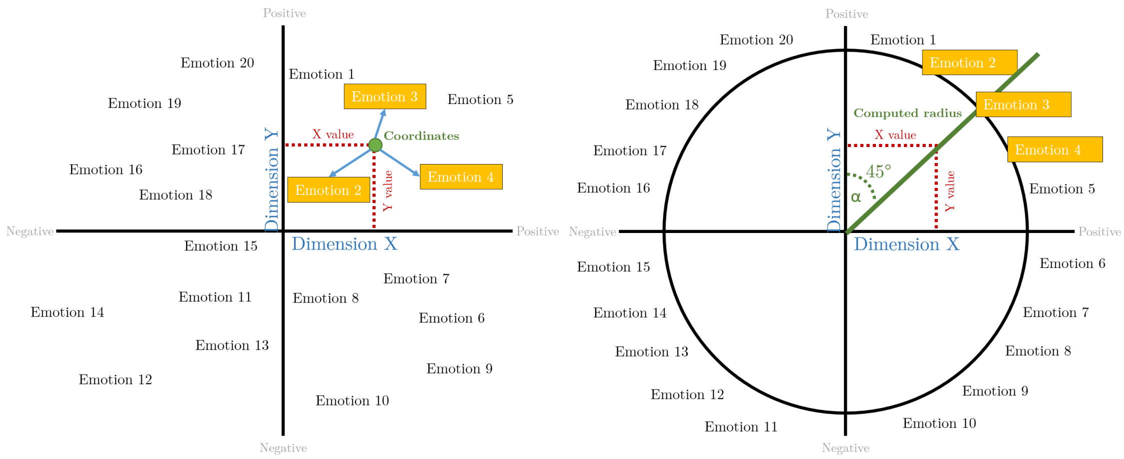


Figure 5.1: Comparison between a two-dimensional affective space for which the *reference frame* is a Cartesian plane (left) or a circumplex (right). The X and Y values corresponds to the E_1 and E_2 values respectively.

5.4.2 The Affective Space as a Circumplex

An alternative for two-dimensional affective spaces consists in the use of a circumplex. A circumplex allows researchers to dispose multiple lexicalized emotions on the edge of a circle, so that the similarity between two items is inversely proportional to their distance on the circumference. The circumplex is determined by an x-axis and a y-axis, usually delimited by opposite poles of the same constructs (*e.g.*, pleasant/unpleasant) or two constructs in antithesis (*e.g.* sadness/enthusiasm). The combination of the two axes creates four quadrants, usually crossing over the negative and positive poles of the two axes. The top-right quadrant is x-positive and y-positive; the bottom-right quadrant is x-positive and y-negative; the bottom-left quadrant is x-negative and y-negative; and the top-left quadrant is x-negative and y-positive.

Given its practical structure, a circumplex can be used to dispose discrete emotions determined by two appraisals criteria in a more *relaxed* way compared to a Cartesian plane. In fact, researchers can tentatively dispose items at an equal distance from each other, assuming the order of disposition on the circle as a sufficient connotation of the affective space. Since the computational model, though, does not use the circumplex as a graphical representation, but rather as a *database* to retrieve suggestions, the fact that the options are not equally spaced on the circumference is irrelevant from a usability standpoint. As a consequence, it is also possible to adopt circumplexes with overlapping lexicalized emotions, or where the distance between any two items on the circle is not necessarily the same, but rather a function of some theoretical or practical mechanism.

On the other hand, the use of a circumplex as the underlying affective space requires a slightly different approach in computing the distance between E and P, because P in a circumplex is best conveyed as an angle, so that $P_{angle} \in \mathbb{R} : 0 \leq P_{angle} < 360$. A possible solution is to use E_1 and E_2 as points in a Cartesian plane, and compute the slope from the origin [0,0] of the plane. The slope can then be used to retrieve the angle from the x-axis, using the inverse tangent, or arctangent trigonometric function, and – with some correction skipped here, see Fritz (2015) or the code repository for more details – also obtain an $E_{angle} \in \mathbb{R} : 0 \leq E_{angle} < 360$:

$$E_{angle} = arctan(\frac{E_1 - 0}{E_2 - 0}) + correction = arctan(\frac{E_1}{E_2}) + correction$$

Once retrieved the E_{angle}, the *k-nearest neighbors* can be determined using the same logic as in a uni-dimensional condition, by computing the absolute distance between E_{angle} and P_{angle}:

$$\Delta(E, P) = |E_{angle} - P_{angle}|$$

Conceptually, with this solution, it is worth considering the case when both E_1 and E_2 are at the neutral point 0, in which case $\Delta(E, P)$ is exactly the same for every subjective feeling, since any point on a circle is at the same distance from the origin[1]. In other words, any subjective feeling is as likely to represent the neutral point as any other. When both E_1 and E_2 are 0, thus, the computational model attributes a virtual random value both to E_1 and E_2 only with the purpose to shuffle the set of subjective feelings and avoid thus that some subjective feelings are arbitrarily suggested more frequently with the neutral point.

Furthermore, the computed slope is a ratio between E_1 and E_2, which entails that some ratios are more frequent than others assuming a sufficiently large range between $[V_{min}{:}V_{max}]$. The more approximate nature of the circumplex disposition is therefore also reflected in a more approximate computation of $\Delta(E, P)$.

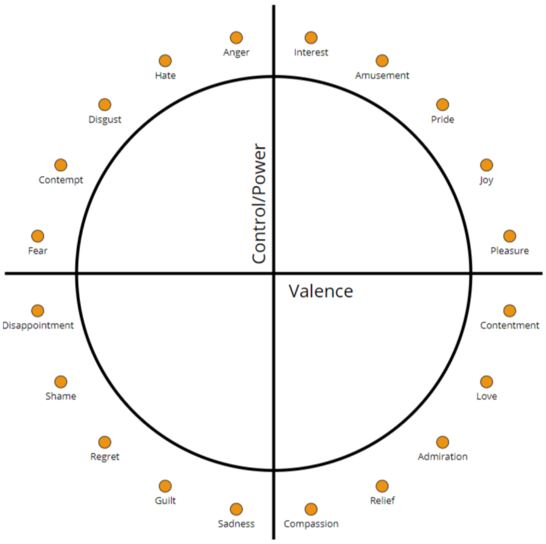
5.4.3 Computational Comparison Between a Cartesian Plane and a Circumplex

Taking advantage of the fact that the disposition of the 20 subjective feelings in the third version of the Geneva Emotion Wheel (Scherer, Shuman, et al., 2013) has been determined using the GRID instrument, the same affective space is available as a circumplex and as a Cartesian plane. As a consequence, it is possible to simulate the behavior of the computational model with the alternative bi-dimensional spaces. For the purpose of the simulation, the following parameters will be adopted:

- The two appraisal dimensions (*Valence* and *Control/Power*) are implemented on a continuum ranging from -100 to 100, with a one-unit step;
- The k-nearest neighbors is set to $k = 3$ and to $k = 10$ to assess how the two models behave when the number of suggested feelings varies;
- The subjective feelings in the circumplex affective space are disposed so that each is equally distant from the previous and the following item, with an initial rotation for the first item of half the distance between each subjective feeling to avoid an option on the 0° – see the right side of Figure 5.2;

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

- The subjective feelings in the Cartesian plan are disposed according to the values retrieved from data available in Scherer, Shuman, et al. (2013), see the left side of Figure 5.2. The original value on a range from -1 to 1 has been multiplied by 100 to map the range of the simulation.



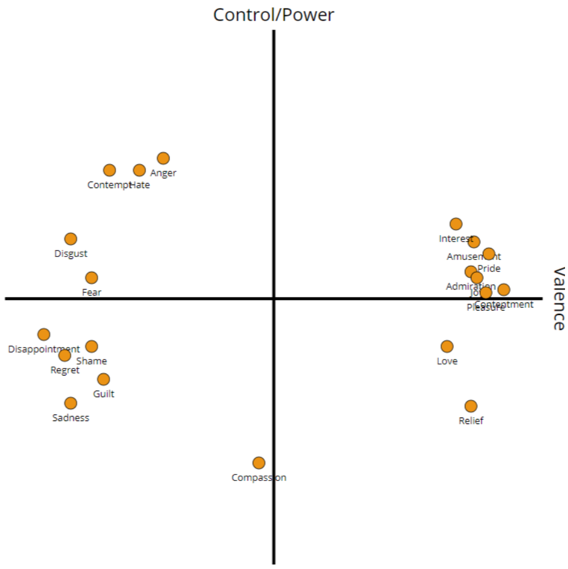


Figure 5.2: Disposition of the Geneva Emotion Wheel's 20 subjective feeling on a circumplex (left) and Cartesian plane (right) affective space. Reconstructed from Scherer, Shuman, et al. (2013).

The simulation is computed by looping the E_1 and E_2 ranges so that each possible combination between the two – that is, [-100, -100], [-100, -99], . . . , [0, 0], . . . , [99, 100], [100, 100] – generates a subset of $k = 3$ or $k = 10$ subjective feelings. Figure 5.3 and 5.3 show the results of the simulation. With a $k = 3$ subset, the radial distance of the circumplex shows a more homogeneous number of appearances of each subjective feeling in the subset compared to the vector distance of the Cartesian plane. The opposite happens with $k = 10$, where the vector distance results in more homogeneous number of appearances compared to the radial distance.

This simulation has only illustrative purpose, because the two affective spaces are only examples of an infinite number of spaces that can be conceived and therefore results may be very different with different configurations. Nevertheless, this kind of simulation may be interesting to carry out before adopting an affective space to assess whether there are subjective feelings that are more likely to be suggested by the system, because that could influence learners in choosing them.

5.5 Three-Dimensional Affective Spaces and Beyond

As previously mentioned, there is growing empirical evidence in suggesting that one- and two-dimensional affective spaces are not enough to account for the variety of emotional experiences in semantic terms (Fontaine, Gillioz, Soriano, & Scherer, 2021; Fontaine et

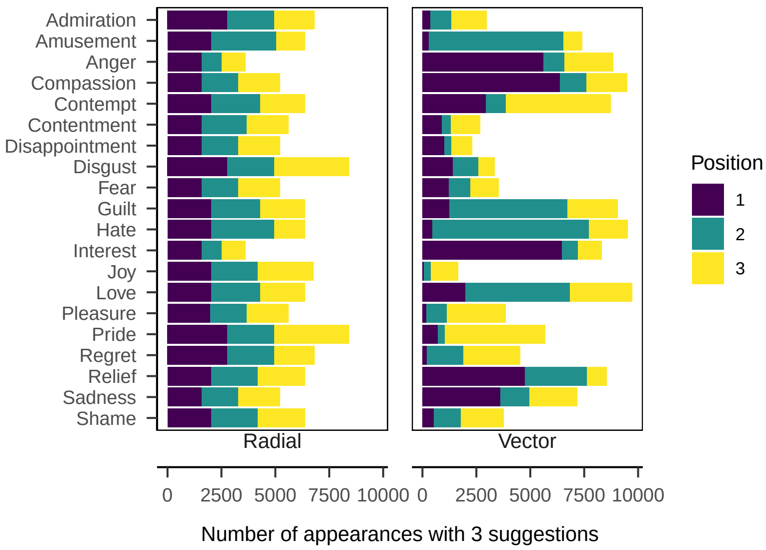


Figure 5.3: Results of the simulation comparing distance computed as a radial or vector distance with a subset of $k = 3$ subjective feelings.

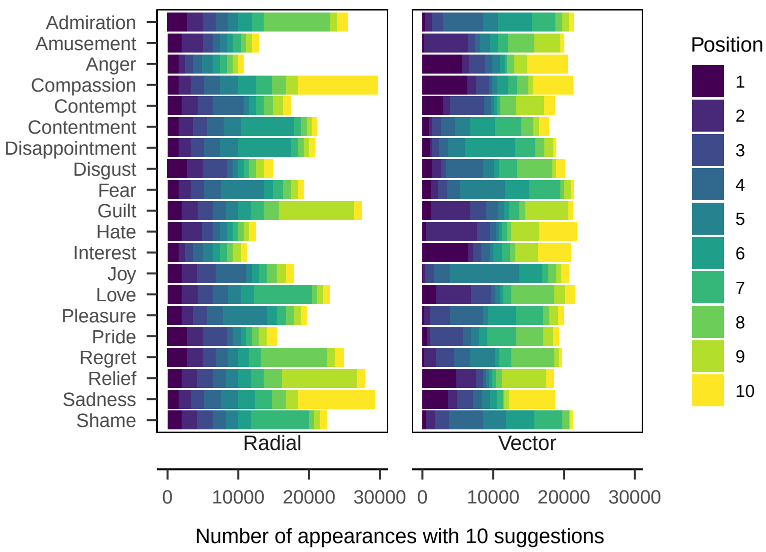


Figure 5.4: Results of the simulation comparing distance computed as a radial or vector distance with a subset of $k = 10$ subjective feelings.

al., 2007; Gillioz et al., 2016). A four dimensional structure determined by *Valence*, *Power*, *Arousal* and *Novelty* seems better fitted for a dimensional representation of emotions (*ibid.*). It is worth mentioning in this regard, though, that Fontaine et al. (2021) found empirical evidence suggesting that the additional *representative* value brought by a third and a fourth dimensions depends on the specific discrete emotions. For instance, terms related to *joy* and *sadness* could be discriminated almost as well with two dimensions as with four. On the other hand, lexicalized emotions closer to *anger* or *surprise* benefited from a four-dimensional structure (Fontaine et al., 2021).

In appraisal theories more specifically, there is virtually no limits to the number of appraisal criteria (Moors et al., 2013). As a reminder, the Component Process Model postulates as many as 13 sequential evaluation checks (3.1). As a consequence, an overarching computational model should be able to integrate higher-order dimensional spaces, where the number of dimensions can be determined by researchers and practitioners with respect to theoretical, practical, or usability criteria (*e.g.*, if the moment-to-moment temporal dimension is particularly important, asking learners to repeatedly rate a high number of criteria would be very time consuming).

Extending the model to more than two dimensions is relatively straightforward when the Euclidean distance is maintained as the algorithm to determine the *k-nearest neighbors*. In fact, it is possible to determine the distance between any two points in a multidimensional space simply by adding additional dimensions to the Pythagorean formula:

$$\Delta(E, P) = \sqrt{(E_1 - P_1)^2 + (E_2 - P_2)^2 + \cdots + (E_n - P_n)^2}$$

The model therefore remains fairly limited in term of computational power required, even if the number of dimensions increases. The difficulty with a higher-order affective spaces is in fact more conceptual than practical, since it is more difficult to imagine a space with three or more dimensions. Beyond three dimensions, it is even difficult to represent the space graphically, which contrasts with the easy-to-understand, and easy-to-implement disposition of subjective feelings in a circumplex. At the same time, the *k-nearest neighbor* approach can precisely contribute to overcome this issue: once a multi-dimensional affective space is established, the dimensions are *flattened* on uni-dimensional ratings. The computational model is then in charge of mapping uni-dimensional ratings into multi-dimensional distances, as vectors between points in spaces the human mind can hardly imagine from a spatial point of view.

As already mentioned, Gillioz and colleagues (2016) adopted the GRID instrument – in its shorter CoreGRID version (Scherer, Fontaine, et al., 2013) – to assess the mapping of 80 emotion terms in an affective space. The number of dimensions and of emotion terms composing the affective space empirically established by the authors is a representative example of how a parsimonious computational model could be deployed in a self-report instrument, allowing users to benefit of all the dimensions and all the suggested emotion terms.

5.6 Limitations Of the Computational Model and Possible Ways to Attenuate Them

The computational model presented here is overtly simple to account for the sheer complexity of *what an emotion is*. It is limited to the appraisal and integration/categorization modules, excluding thus three out of five components of the Component Process Model. The complexity of the two modules is also simplified. In absolute terms, thus, the model reflects only a fraction of the theoretical work of the CPM, appraisal theories, or emotion theories more broadly. Compared to other existing solutions, though, the model adds some distinctive features. The limitations of the tool are therefore assessed with respect to these distinctive features.

A first limitation consists in the need for the model to dispose of an accurate *reference frame*. Compared to a pure dimensional approach, a pure discrete emotion approach, or a combined – but not integrated – approach, the model highly depends on the specific details of the underlying affective space. As a reminder, for instance, the EMORE-L (Gaëlle Molinari et al., 2016) combines dimensions and discrete emotions, but without a link between the two. The *emot-control* (Feidakis et al., 2014; Feidakis et al., 2013) and Mood Meter application (Brackett et al., 2019; Hoffmann et al., 2020; Nathanson et al., 2016), as well as the Geneva Emotion Wheel, combine the two but only from a dispositional point of view, with discrete emotion that are nicely *spread* over the space to an equal distance for graphical ease of use. With the computational model at hand, the precise coordinates of each discrete emotion on the underlying affective space is taken into account *as is*. The simulations in Section 5.4.3 have already highlighted this limitation. When the lexicalized emotion do not *occupy* the whole range from V_{min} to V_{max}, there are discrete emotions that will appear more often in the subset, especially the ones at the edge of isolated clusters. Those will still be considered *nearest neighbors* when E is actually in uncharted (*i.e.*, blank) territory.

This constraints impacts the model in another way. In fact, the model subsumes a linear relationship between the dimensions and the corresponding subjective feeling. Using statistical jargon widely adopted in psychology, the model does not take into account the possibility of *interactions*, whereas some subjective feelings may in fact be more likely to occur given a dimension conditional to the value of the other (Fontaine et al., 2021). For instance, a subjective feeling on a dimension may be more probable when the evaluation of another dimension is either very high or very low, but not *neutral*. On the contrary, the model expects the subjective feeling to be monotonically (*i.e.*, proportionally) linked to both dimensions. Fontaine et al. (2021) have tested this possibility on the four-dimensional structure already depicted. They indeed found some interactions, especially when the third and fourth dimension are stacked in.

Given the intrinsic simplicity of the computational model, it was inevitable that some limitations occur. Any computational model must be calibrated around a ratio between the quality/reliability of the information *in* and the quality/reliability of the information *out*. (McElreath, 2020). The flexibility of the computational model, though, can also be exploited to attenuate the limitations. Existing affective spaces are mainly driven from the theoretical perspective, meaning that they strive for accuracy. Depending on researchers and practitioners' specific aims, one can contemplate the possibility to apply a function $f(P)$ that transforms the affective space, for instance by *distancing* the discrete emotions as much as possible. The f applied may be linear, but also non-linear, which may be useful for instance if one wants to merge existing affective dimensions into composite criteria (Fontaine et al., 2021). This possibilities are nevertheless not explored in the present contribution.

Other ways to attenuate the limitations consists in adapting the interface of the DEW. As seen in Section 4.3.2, the tool can also integrate an "Other. . . " field, through which respondents can provide their own subjective feeling. With this option, the computational model is by-passed by respondents own interpretation of which lexicalized emotion represents at best their own subjective feeling.

5.7 Summary

This chapter illustrated a parsimonious computational model at the core of the Dynamic Emotion Wheel. The model takes an n-dimensional affective space as an underlying *reference frame*, and compares the actual evaluation made by users with the predicted value of the pre-defined discrete emotion in the affective space. The model then computes, using a k-nearest neighbors approach, the discrete emotion who minimize the distance between the actual evaluation and the predicted value. These become the k suggestions that the model propose to the user as the most likely subjective feeling corresponding to the given dimensional evaluation. The computational model is therefore a form of *helper* or affective assistant, which gives *informed suggestions*, but that also reckons its limits. As a consequence, it is all the more important that the interface of the tool may be determined as much as possible from researchers and practitioners themselves, which is the subject of the next chapter.

Chapter 6

Building a Toolbox Around the Parsimonious Computational Model

The previous chapter illustrated the parsimonious computational model linking appraisal dimensions and subjective feelings, which is the *core* of the proof of concept proposed by the thesis. This chapter introduces a toolbox built around the computational model, which allows the configuration of an Emotion Awareness Tool (EAT) according to researchers and practitioners needs. As stated in the introduction, the implementation of a toolbox is driven by open science tenets, which advocate the need for investing in the development of open material that can foster transparency, sharing, comparison, replicability and reproducibility of research practices, measures and results (Chambers, 2017; Flake & Fried, 2020; Lowndes et al., 2017; L. D. Nelson, Simmons, & Simonsohn, 2018; Nosek et al., 2015; Scheel et al., 2020a; Scherer, 2005). The chapter starts by a general overview of the toolbox, which is divided in an area dedicated to researchers and practitioners, and the *end-user* interface of the tool integrated in a computer-mediated learning environment. The reminder of the chapter then illustrates the available configuration options of how emotion is expressed-displayed and how emotion is perceived-monitored respectively. Finally, the chapter describes some features of the toolbox that can be configured from an empirical point of view, as well as the measures that can be retrieved from the toolbox.

6.1 General Presentation of the Toolbox

The toolbox is a web-based application that can also be installed on local machines in relatively few steps. The details are ignored here because this sort of information is likely to change over time, but the choice of a web-based application has been made to maximize ease of access to both researchers and users, with only a modern browser as a requirement. In fact, contrary to *pure* and individual self-report instruments, which may be administered in different settings with minimal adaptations, an EAT presents more challenging technical constraints, especially from a moment-to-moment and inter-personal perspective. Whenever synchronization between two or more *agents* sharing information is concerned, it is necessary to adopt a bi-directional communication technique, such as web sockets. Most web applications are in fact based on the client-server architecture: the *agent* (the client) sends a request to the server, which responds back with the appropriate response (e.g., the content of a page). Outside the bounded episode of request/response, though, the server has no way to send information to the client, if the client hasn't initialized a request first. Nowadays the concept of *push* notifications in mobile phones or web sites has become ubiquitous, but still, the technique requires adequate software, which is not readily available. The combination of theoretical/pedagogical and technical requirements, thus, advocate the development of a comprehensive toolbox allowing researcher to customize and carry out directly their instances of the EAT according to their specific aims, minimizing the technical burden to set up a computer-mediated environment. This section provides a general overview of the toolbox, which comprises a *front-end* part dedicated to users, and a *back-end* part dedicated to researchers and practitioners. The section starts with the overview of the *back-end* part, since it is the one where the *front-end* part is defined.

6.1.1 The Toolbox Back-End

The toolbox back-end is similar to a resitrcted admin area, where researchers must provide their credentials to access their instances of the tool. Figure 6.1 shows the dashboard of the admin area available once the researcher has logged-in. At the time of writing, the back-end is available only in English.

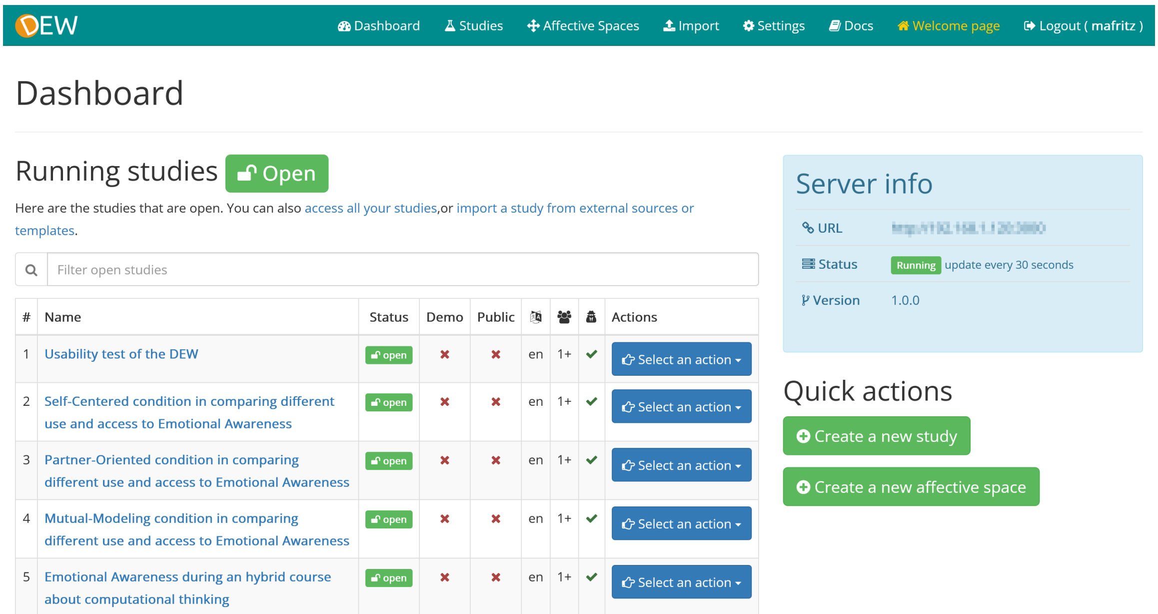


Figure 6.1: The dashboard page of the admin area shows a list of active instances/studies and gives access to the different features of the back-end.

From the back-end part of the toolbox, researchers and practitioners can perform mainly four actions that have scientific and/or pedagogical relevance :

1. Access instances – that is, specific configurations of the tool – which are identified as *studies* in the interface[1]. The dashboard only lists the active instances, but in the *studies* tab researchers can also find past or inactive instances.
2. Manage instances by creating, duplicating, modifying, removing or exporting *studies*.
3. Manage affective spaces that can be used in *studies*. Affective spaces can be either private, available thus only to the logged-in researcher, or shared with other researchers subscribed to the same server.
4. Import configurations of *studies* or affective spaces from external sources, for instance – ideally – supplementary material from a contribution that has used the toolbox.

A *study* is the element of the toolbox that glues all parts together. It is determined by the different configurations options and is made available to users through the toolbox front-end illustrated in the next subsection. Figure 6.2 shows a cropped detail page of a *study*, which happens to be an archived instance of the usability test illustrated in Chapter 4.

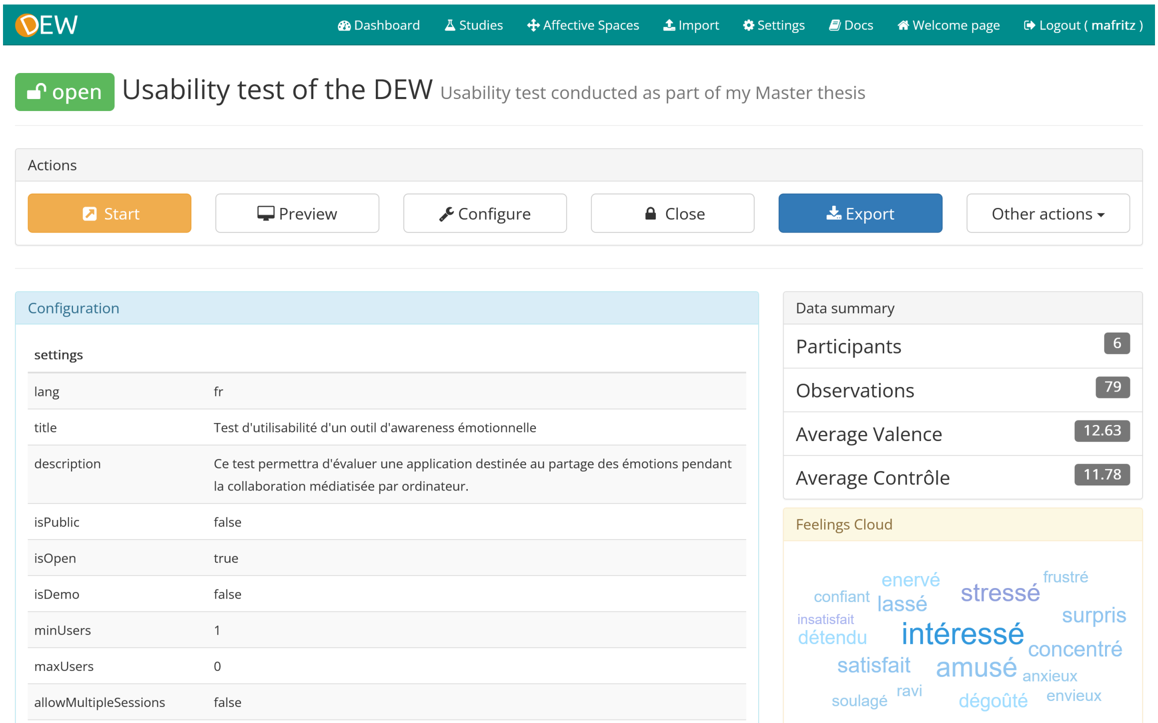


Figure 6.2: Example of a *study*'s detail page. The screen capture has been cropped to save space.

Through the detail page, it is possible to perform a series of actions connected to the *study*. Only the most salient actions are listed hereafter.

First, the study can be started, which means that the front-end of the instance is launched in a new tab of the window. Each study disposes of its own web address, which

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

can be limited to the local machine or local network, or publicly available over the web depending on where the server is installed. This is thus just one way to start-up an instance, which can nevertheless be accessed directly by users through the URL, without the need to pass from the admin area. The study can also be accessed in a preview mode, which shows how the final instance will look like to user, but without gathering any data.

Second, the *study* can be configured. The configuration is equivalent when creating a new instance or modifying an existing one. Instances can also be duplicated, which may be useful in case of conditions within the same *study*. The configuration details are illustrated in the following sections of the chapter.

Third, the status of the study can be easily updated, switching from *open* to *closed* or vice-versa. It is in fact possible to decide whether an instance is available or not at any given time. When not available, the URL of the front-end will show a message informing the user that the instance is not active.

Fourth, it is possible to export different elements of the study, as depicted in Figure 6.3. The first and more obvious element to export are the data collected. In this regard, the detail page also provides a data summary frame on the right-side of the screen, right under the main action buttons. The data summary report the number of participants who have been used the instance, the number of observations/emotions expressed, as well as an average for every dimension rated. The recorded observations are also illustrated in a word cloud, which provides a perceptive representation of the most frequent subjective feelings used by participants. A second element that can be exported are the participants accesses to the instance. Since it is possible that a participant access the instance without expressing any emotion, this form of data can still be used as a separate dataset. A third element concerns the overall configuration of the *study*, which can be shared with other interested people, for instance as supplementary material in a scientific contribution. Finally, the whole *study* (data, affective space, and overall configuration combined) can be exported as a form of backup.

From the admin area it is also possible to access the list of available affective spaces, as well as create new ones. As stated above, affective spaces can be either private or shared. Only the researcher/practitioner that has created or added the affective space to the server can decide to make it public or private. When used in a *study*, the affective space is nevertheless copied into the specific instance, so that actions performed to the affective space do not automatically propagate to *studies*.

Alongside the information necessary to represent the structure of the affective space (*i.e.*, dimensions and disposition of lexicalized emotions), the toolbox also reckons the importance of crediting the original author(s) of the affective spaces. In this regard, it is possible to specify how each affective space should be cited. Figure 6.4 shows the detail page of the English version of EATMINT circumplex introduced in Section 4.3.3. The structure of the affective space can be exported either as data, with the *Export* button on the top-left side of the screen, or – for one- and two-dimensional spaces only – a Scalable Vector Graphic (SVG), that is an image that can later be easily modified by a vector editor[2]. The SVG file

[2]At the time of writing, examples of widely adopted vector editors are the proprietary Adobe Illustrator and Affinity Designer, or the open-source Inkscape.

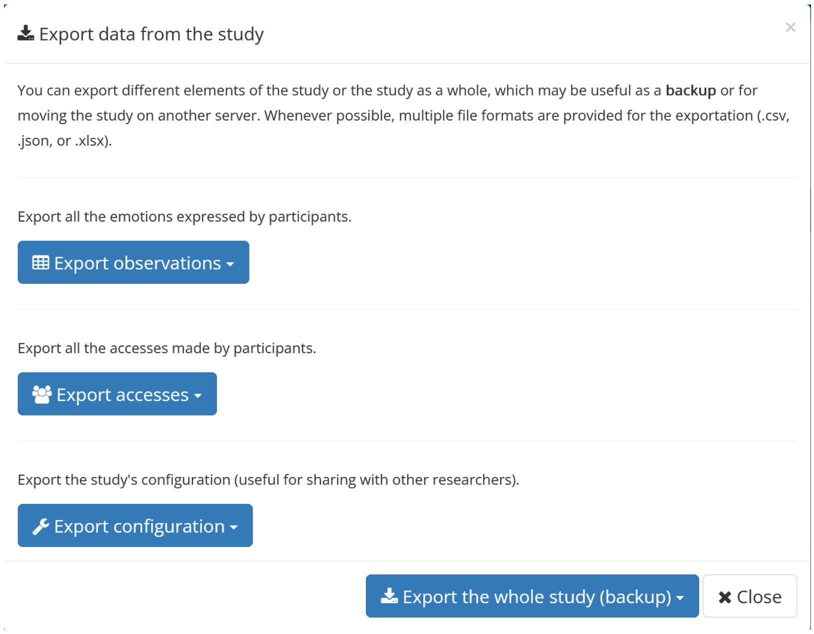


Figure 6.3: Screen capture of the export procedure, which can be executed on different elements of the *study* and, if possible, with different file types.

can later be included directly, or rasterized before inclusion in a contribution for illustrative purposes.

At the time of writing, there are several affective spaces identified in the literature that can be implemented as underlying structures for the DEW. These include the EATMINT circumplex in French and English Fritz & Bétrancourt (2017), the three versions of the Geneva Emotion Wheel Scherer, Shuman, et al. (2013), the circular scaling coordinates of 28 affect words by James A. Russell (1980), and the mapping of 80 emotion terms on a four-dimensional space by Gillioz et al. (2016). Furthermore, as stated in the description of the computational model in the previous chapter, an higher-order dimensional space can also be used for lower-order rating dimensions. For instance, a three-dimensional affective space can be (1) disjointed in three different one-dimensional affective space, (2) combined in two different bi-dimensional affective space, or (3) used in its whole three-dimensional structure. Bi-dimensional affective spaces can also be represented both as a circumplex or as a Cartesian plane, allowing thus greater flexibility on the underlying *reference frame* to be adopted.

Finally, the last relevant feature from a scientific or pedagogical perspective concerns the possibility to import affective spaces and whole *studies*' configurations. As mentioned above, this is meant to increase the transparency and sharing of research and practices involving the use of the tool. External sources can be imported for direct use, for better understanding of the inner functioning of an instance, as well as for checking the accuracy and adherence to protocols.

To sum up, the back-end area of the toolbox is meant to facilitate researchers and practitioners tasks both at the individual and collective levels. At the same time, there is also the possibility to use the tool in isolation, on individual premises, with relatively few

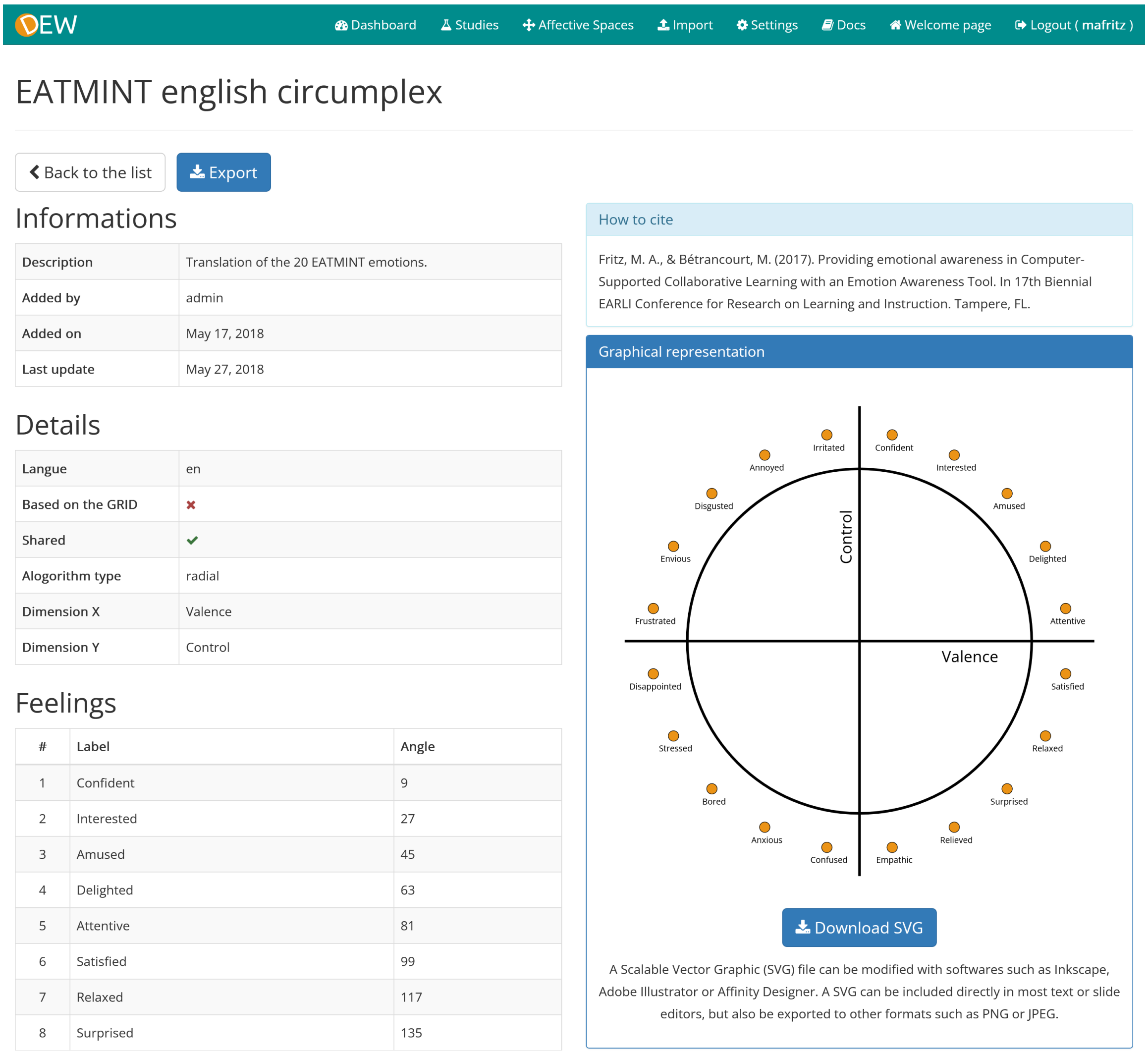


Figure 6.4: Detail page of the English version of the EATMINT circumplex presented in Section 4.3.3 as an example of affective space. Each affective space disposes of a field to indicate how it should be cited in order to credit the author(s).

technical steps and requirements.

6.1.2 The Toolbox Front-End

The toolbox front-end is the instance of the tool participants/learners will interact with. It is available through an URL as any other site or web application. At the time of writing, the *fixed* parts of the front-end interface – which are few considering that most of it can be configured – are available in English, French, German and Italian.

The front-end part of the screen can count up to four different screens: (1) one for general instructions, (2) one for the log-in/synchronization with two or more users if needed, (3) one for the task itself, and (4) one for a debriefing or end-of-task screen that can for instance provide a link to further steps in the procedure. There are nevertheless ways to modify the

URL given to users, for them to skip any step except the task itself. For instance, users can be provided with an URL that includes a specific or random ID and that logs users directly into the task. This mechanism has been used in the empirical contribution of Chapter 8.

When two or more participants have to be synchronized to start a specific task, the toolbox provide a *pseudo* log-in system that can be configured to hold-on until all the required users are ready to start the task. Figure 6.5 illustrates this passage. The same instance can be used for as many groups as needed, with the size of the group that can potentially scale up to thousands. The limit is in fact determined by the machine's ability to deal with concurrent sockets. Users belonging to the same group need only to provide a common Group ID. Once again, the instructions on how to log-in into the system can be customized. Synchronization is also not mandatory. For instance, in a longitudinal use of the EAT as the one that will be made in the second empirical contribution of the thesis, participants could log-in at any time, regardless of the number of concurrent colleagues connected. The system automatically detects if a new (or returning) user has accessed the instance and adapts the interface accordingly.

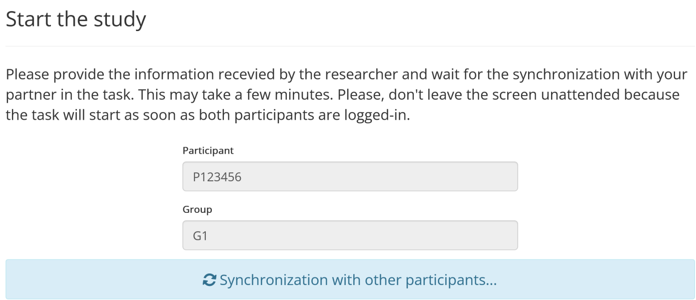


Figure 6.5: Example of *log-in* screen when the task has to start at the same time for all involved users. This is nevertheless not a mandatory step.

The core of the front-end is the interface where the EAT is available, which can be coupled (or not) with a task or learning activity on the side. Figure 6.6 shows the front-end interface of an hypothetical study or learning activity based on collaborative writing. The EAT occupies the left-hand side of the screen, whereas the online editor is in a iframe (technically a window inside a window) on the right-hand side. The interface may also presents a timer and a button that allows participants to stop the task at any time. This is just an example of how the choice of a web-based application allows the integration with existing tools, but once again, the possibilities are much wider and depends on the specific aims for which the toolbox is adopted. As it will be described in the reminder of the chapter, many aspects of the overall interface and the inner functions can be adapted through the back-end part of the toolbox. The example, for instance, adopts an *Energy* x *Familiarity* dimensions, which are derived from the *Arousal* x *Novelty* dimensions of the four-dimensional structure in Gillioz et al. (2016).

It is important to note that the toolbox guarantees only the synchronization between the emotional information shared by the EAT, as well as the beginning and the end of the activity if these parameters are specified. The presence of a coordinate task must be taken into account by the specific tool adopted (*e.g.*, an online text editor). In this regard, the

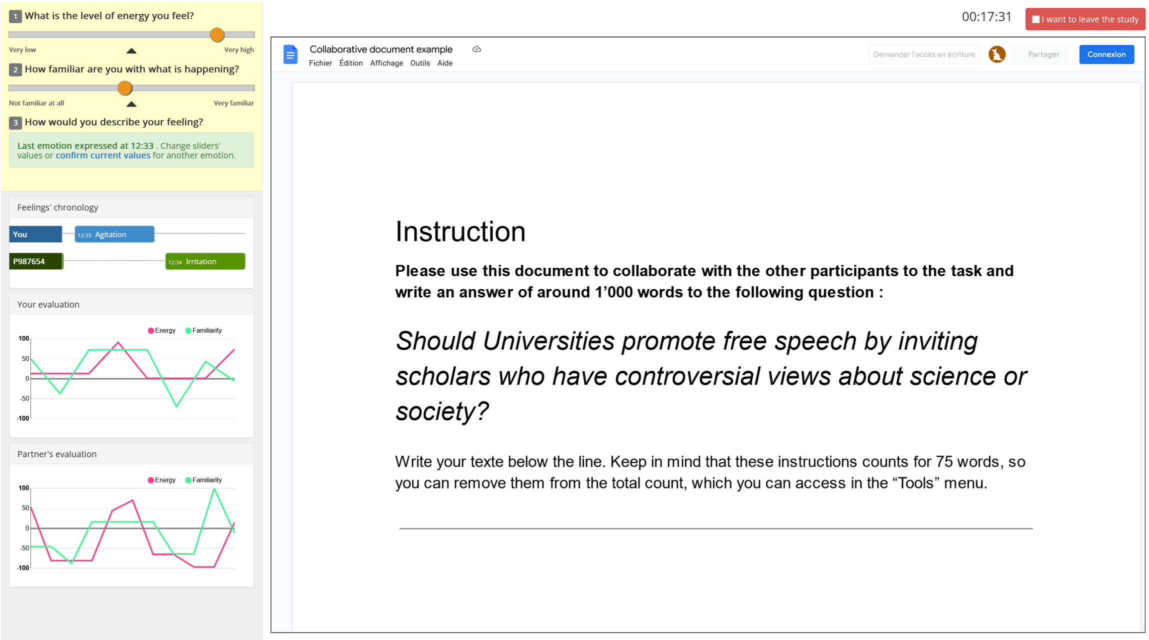


Figure 6.6: Example of an instance of the toolbox comprising the EAT on the left-hand side and a collaborative writing task on the right-hand side.

overall toolbox passes to the frame within which the task is performed the identity of the participant and the group, as well as the emotion shared through the interface. Researchers and practitioners disposing of the technical know-how to intercept that information can thus use it.

The synchronization between the interfaces of users is made through the bi-directional technology mentioned above, which guarantees the advocated moment-to-moment feature of the EAT. The speed of the bi-directional communication may vary slightly depending on how the clients and the server are connected (e.g. on a local network or over the web), but the latency remains in the order of milliseconds.

To sum-up, the front-end of the toolbox consists in a web-application of the sort that most users should be nowadays familiar with. It guarantees bi-directional communication over very brief periods of time for the emotional information, as well as the onset and end of the activity if needed. The overall interface reflects the choices made by the researcher or practitioner when the instance was set up through the back-end. Some of the most salient choices are briefly outlines in the reminder of the chapter.

6.2 Configuration of the Expressing-Displaying Function

The options available about how emotion is *encoded* into the system pertains to two intertwined element: the underlying affective space and how the expressing-displaying part of the EAT is configured. The options available for the affective space have already been outlined in the general presentation of the back-end part of the toolbox in Section 6.1.1. It is worth reminding to this purpose that any kind of affective space can be adopted, as long as it meets the requirements of the underlying computational model defined in Section 5.2.3. Existing and validated affective spaces may be preferable in some circumstances, whereas more exploratory and custom choices could be warranted in others. The fact that affective spaces can be shared within the same server and across servers if needed is precisely meant to leave the maximum freedom, but at the same time foster transparency and comparability of conditions. An affective space is technically determined by the following minimal information:

1. The number and labels of the dimensions, from a minimum of 1 to potentially no upper limit, even though a large number of dimensions may be unfit for the interface.

2. The type of algorithm adopted to determine the *k-nearest neighbors*. This depends on the structure of the underlying affective space:

   - For uni-dimensional spaces, as well as for three-dimensional spaces and beyond, the available algorithms are the vector/Euclidean distance and a random order, which may be interesting for research purposes (see Section below). The random order is based upon the Fisher-Yates algorithm[3] to shuffle an array of elements.
   - For two-dimensional spaces, the available algorithms are the ones listed above, plus the radial/arctangent distance.

3. A non-empty array of lexicalized emotions, each of them with their position relative to the underlying affective space and the specific label. The position can be encoded as a number of coordinates or as an angle for the bi-dimensional circumplex. At the time of writing, the toolbox adopts a system of coordinates ranging from -100 to 100 with a one-step increment, but this may become more flexible in the future, for instance to account for affective spaces derived from theories with a fixed number of values (*e.g.* Pekrun's Control-Value theory).

The expressing-displaying interface of the tool can be modified almost on every aspects. Researchers and practitioners can define:

1. The specific way by which each the rating of each dimension is prompted (*e.g.*, *Is the situation pleasant?*, *How familiar are you with the situation?*, etc.), as well as the labels of the two opposing poles of the continuum. Those can be general (*e.g., Not at all* $\leftrightarrow$ *Yes, absolutely*) or more specifics to the dimension(*e.g.*, *Very unfamiliar* $\leftrightarrow$ *Totally familiar*).

2. How the choice of the lexicalized emotion is prompted (*e.g. How do you feel?*, *What emotion do you feel?, ...*).

3. The number of lexicalized emotions that are proposed as buttons on the interface, with any number less or equal to the number of subjective feelings in the affective space. A greater number of options means that more buttons populate the interface, even though it is worth noting that the interface stretches when the buttons are

[3] https://en.wikipedia.org/wiki/Fisher%E2%80%93Yates_shuffle

available, but than shrinks again when the choice has been made. As a consequence, a large number of options would not reduce permanently the space available for the perceiving-monitoring function. Conversely, this number may very well be 0, meaning that no suggestion is made at all.

4. Whether to add the "Other…" text-field to the interface or not. Adding it is generally considered a good practice (Mortillaro & Mehu, 2015; Scherer, 2005), but there may also be situations in which a forced choice can be of interest.

5. Whether to add or not an alternative button of the kind *No emotion* or *Not sure*. The precise label of this additional button can also be determined.

The customization of how emotion is encoded can be fully appreciated by using *extreme* cases. The simplest case consists in a *pure* dimensional approach. In fact, it is possible to by-pass the computational model by providing an empty affective space, or by setting the number of suggested emotion to 0. In this case, the alternative button can be labelled the like of *Send*, and an emotion thus be represented only by the values on the sliders. From a theoretical perspective, this combination would be closer to dimensional/constructivist theories of emotion (James A. Russell, 2003, 2009; Stanley & Meyer, 2009) – se also Section 3.2 – especially if the alternative button *Send* is replaced by an open text-field, allowing respondents to construe the appropriate emotional label based on their own cultural background and vocabulary.

Another extreme case consists in setting the number of suggestion at the same value as the total number of subjective feelings in the underlying affective space. That would by-bass the sub-setting, but not the sorting part of the computational model. In fact, all the lexicalized emotion would be available on screen, but their precise order would depend on $\Delta(E, P)$. This would be more aligned with those appraisal theories of emotions that subsume the existence of representative/modal emotion families (Grandjean et al., 2008; Moors et al., 2013).

It is also possible to provide the very same coordinates to all the subjective feelings in the affective space, in which case the order of the buttons will be fixed, no matter the ratings on the sliders. This may for instance allow to replicate part of the EMORE-L interface (Gaëlle Molinari et al., 2016), where the three appraisal dimensions and the eight discrete emotions are not linked. One can imagine to contrast a condition in which there is no link between the dimensions and the discrete emotions, and one when the link is done through an appropriate underlying affective space. As a reminder from the related works in Section 2.2, the authors found that discrete emotions that are theoretically distinguished by different appraisal profiles (*i.e.*, different P), received by contrast similar ratings on the *Control*, *Value* and *Activation* dimensions (*i.e.*, similar E). The different configuration of the EAT could therefore investigate whether the computational model guides learners to have a more *adequate* (Scherer, 2007) appraisal evaluation, as illustrated in Section 3.5.1.

Finally, users landing on the front-end interface are not aware that they are entering a platform for emotional awareness. As a consequence, the interface can be configured without any reference to emotion. One may set up a *non-affective space* in which the dimensions are, for instance, evaluation of the quality of the colleague's work, or information closer to a

Knowledge Awareness Tool (Engelmann et al., 2009; Gaëlle Molinari et al., 2009; Ogata & Yano, 2000) rather than an Emotional Awareness Tool. As long as meaningful information can be conveyed through dimensions and discrete elements, the same environment can be configured with different objectives.

To sum-up, the available configuration options in the expressing-displaying function can have pedagogical, theoretical, as well as user-experience implications. The flexibility of the configuration bestow researchers and practitioners with the choice of which aspect should be privileged.

6.3 Configuration of the Perceiving-Monitoring Function

As already mentioned, whereas the way emotion is *encoded* may be kept more or less the same regardless of the context, emotion *decoding* is highly dependent on the purpose of emotional awareness and how data is graphically represented from a visuo-spatial perspective (Hegarty, 2011). As illustrated in the abstract model in Section 2.5, a multi-purpose proof of concept shall consider emotional awareness not only based on inter-personal, or not only limited to dydic interactions. Computer-mediated learning enviroments in particular span from the individual to the thousands of learners of MOOCs. Ideally, a multipurpose EAT should be able to scale congruently to the context. This possibility is inherently linked with the perceiving-monitoring part of the tool. The two representations (emotion timeline and appraisal line charts) proposed by the prototype are clearly limited to a restricted number of users. Adding 20 rows of the timeline and 20 line charts to account for 20 learners would pile up an overwhelming amount of information on screen. Researchers and practitioners may therefore also be allowed to choose between graphical representations of emotion aligned with the number of learners sharing the same EAT, as well as the kind of emotional meaning-making that the tool is supposed to foster (Berset, 2018; Derick et al., 2017; Fritz, 2016b; Leony et al., 2013).

This subject, though, has received so far limited attention overall, as mentioned in Section 2.4. The work on the proof of concept has unfortunately not progressed very much on this topic either, in spite of the good intentions (Berset, 2018; Fritz, 2016b). In particular, Berset (2018), in his Master thesis, attempted to conceptualize graphical representation of emotion directly linked to the Dynamic Emotion Wheel, which were then assessed in a usability test. At that time, the DEW was still a prototype limited to two dimensions, and therefore the author focused on representations that combined two dimensional values with a discrete emotion. In spite of the difficult task, Berset (*ibid.*) proposed interesting representations – such as timeline that *piles* up discrete emotions by showing also their frequency (see Figure 21 in the original manuscript) – that have unfortunately not yet been implemented into the tool.

As a consequence, the perceiving-monitoring function of the toolbox is currently limited. This sections nevertheless provides an overview of the available configuration options.

The first option, which has implications regardless of the specific graphical representation adopted, is relative to the persistence of data over time. In this regard, it is possible to decide whether the expressed emotions are maintained between different *sessions* with the instance or not. A session corresponds to a new access to the tool, which is the equivalent to loading the page into the browser and *log-in* into the instance. When data is not persisted, all the graphical representation restart afresh. If persistence is maintained, the users will find all the available data from the first time the instance has been adopted. This also includes, in case of longitudinal inter-personal/group-awareness, data produced by other users sharing the same group in the instance.

At the time of writing, the toolbox proposes only three kind of graphical representation of emotions, depicted side-by-side in Figure 6.7: (1) the emotion timeline, (2) the dimensional/appraisal line charts, and (3) a word-cloud similar to the one available in the *study*'s detail page in the back-end part illustrated in Figure 6.2. All the representations will be adopted in the empirical contribution of Part III, so they are presented here only briefly.

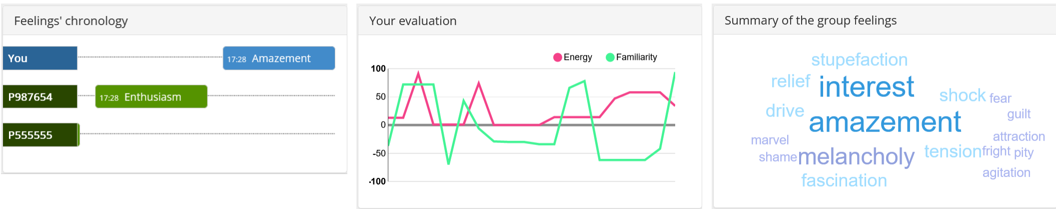


Figure 6.7: Three graphical representation of the perceiving-monitoring function: the emotion timeline, the appraisal/dimensional line charts, and the subjective feelings word-cloud. Each of them can be configured.

The three graphical representation can all be configured according to the following parameters. Congruently with the intra-personal and/or intra-personal function of emotional awareness, each of them can provide information about the user herself and/or the other users in the same group/instance:

1. For the emotion timeline, this means that the widget can show (1) only the *You* row; (2) only one row for every other user (*e.g.*, the *P987654* and *P555555* rows), whereas it is not possible to single-out exclusively one user if many are sharing the same space; or (3) both the *You* and one row for each other user, as it is the case in the image.
2. For the line charts, the interface can show (1) only the line chart with the appraisals/dimensions of the user herself (as in the image); (2) only one line chart for every other user in the group; or (3) both the user's own line chart and that/those of the other(s). The legend of each line/dimension can also be configured.
3. For the word-cloud the situation is slightly different, but the sense is the same. The interface can once again show (1) only the cloud of the user's own feelings; (2) only one cloud where the feelings of all other users (bar the user herself) are combined; (3) only one cloud where all the feelings (user herself plus all other users) are combined; or (4) any combination of the previous three.

The title of the panel in which the graphical representation is contained can also be configured. In the case the line charts of other participants are adopted when there are more than one other user, the users' ID is added to the panel's title. The order of appearance can also be determined thorough the back-end admin area, except within a sequence of line charts belonging to other users, which will be stacked one after the other.

To sum-up, at the time of writing there is a limited set of graphical representation to *decode* emotion for the perceiving-monitoring function of awareness. Those that are available, though, can be articulated from an intra-personal, inter-personal or combined perspective. This possibility will be harnessed in two empirical contributions in Part III.

6.4 Configuration of Research-Related Features

Being the toolbox also meant for researchers, there are a few features that are specifically meant for empirical studies or experiments. Most of them belong to standard procedures in studies' interactions, whereas one is very specific to the toolbox. The *standard* features are the following:

1. As mentioned in the presentation of toolbox front-end, a screen can be devoted to instructions. Even though this can be helpful also for practitioners, they may be particularly useful in case of online studies. At the time of writing, the text can be written in Markdown, a widely used, light-weight mark-up language that allows to structure a text with titles, lists, images, etc. So the instruction can be as long and as precise as needed.

2. It is possible to force participants to use a pre-defined or random IDs for either or both the user and group identifiers. The system can also block the same IDs from being used again once the log-in has been done.

3. The duration of the task can be set to a specific time, with or without a timer that is available on the top of the interface, as in Figure 6.6. A button that users can click to stop the task can also be made available, and labelled as wished.

4. The last screen of the front-end interface can contain a custom text, which can also be formatted. It is also possible to provide an *exit URL* and a label for the button that points at it. This may be useful for instance if there is an additional step in the procedure (*e.g.*, an online survey) or for directing participants to the end-of-trial URL commonly adopted by online recruiting platforms.

The specific research-related features consists in the possibility to simulate the emotional experience of one or more fake participants, and inject it in the interface through the graphical representations of the perceiving-monitoring function described above. This possibility has been evoked in the illustration of the DEW's usability test and will also be adopted in the first empirical contribution presented in Chapter 7. To simulate the emotional experience, researchers must provide a list of *observations* that respects the number of dimensions of the underlying affective space, for consistency with what is expressed by *true* participants. The list of simulated emotions is therefore composed by:

1. A consistent identifier for every emotion belonging to the fake user. For instance, in the emotion timeline in Figure 6.7, the *P555555* is simulated. This shows incidentally that a fake participant can be added as the only other participant to a dyadic task, but also in addition to more than one *true* participant. To simulate more than one fake user, it suffices to enter in the list a different number of unique identifiers.

2. A time when the simulated emotion will be injected into the system. The value refers to the time elapsed since the log-in from every *true* participant accessing the instance, so it is more appropriate to use the simulation with synchronized trials of a fixed duration and not repeatable more than one time.

3. A value respecting the domain $[V_{min} : V\_max]$ for every appraisal/dimension in the adopted affective space. If the interface includes the line charts for other participants, those values will serve to populate the graphics.

4. An emotion term or idiom, such a lexicalized emotion, that represents the subjective feeling. The label does not necessarily have to be part of the adopted underlying affective space. It can be manipulated as desired. For instance, one can test whether a fake participant that only uses custom emotion terms influences *true* participants to avoid the pre-defined options. Or the fake participant could provide lexicalized emotions belonging to the affective space, but with incoherent appraisal/dimensional values (*e.g.*, a negatively valenced emotion terms when the *Valence* dimension is rated positively, or vice-versa).

The simulation of a participant can also be adopted systematically. As a reminder from the related works in Section 2.3 about inter-personal emotional awareness, Avry, Chanel, et al. (2020) manipulated feedback in an online collaborative game representing the *Control* and *Value* dimensions of Pekrun's theory (Pekrun, 2006; Pekrun & Perry, 2014). The toolbox can provide a simulation including also the subjective feeling rather than only appraisal dimensions. For instance, different *emotional profiles* of the *fake* participants can vary according to how it evaluates and categorize/integrate the situation. The consequences of these differences can be tested on a number of cues spanning the quality of interaction, what kind of personality is attributed to the *fake* participant, whether it impacts the emotional experience of the *true* participants, and so on. In this regard, a third experiment harnessing this features was planned for this contribution, but was ultimately not carried out do to the outbreak of the COVID19 pandemic.

To sum up, the toolbox provides researchers with features that can serve *standard* procedures in empirical settings, but also use the tool to create conditions or manipulations that can help to respond to specific research questions. Measurements gathered through the toolbox play in this sense a major role and are thus presented next.

6.5 Measures Available Through the Toolbox

Whereas the features illustrated in the previous section are mainly meant for researchers, the measures available through the toolbox may also be of interest for practitioners to have a sense of the *emotional experience* of learners. As illustrated in the general presentation of the toolbox back-end, the measures can be exported from the admin area in different file formats, which are meant to ease subsequent analysis, but also sharing in open-format (Gilmore et al., 2018; Levenstein & Lyle, 2018). Data available through the toolbox can be divided in accesses and observations.

6.5.1 Data About Accesses

Accesses consists in initializing a new session with the instance of the toolbox by visiting the *study*'s URL. Given the web-based nature of the toolbox, a session is consistent with a visit of a web page, which technically consists in a new HTTP request sent from the client to the server. This also includes, thus, page refresh in the browser. If the browser's window or tab remains open, though, the connection is maintained even if the window or tab are unattended by the user. A new access is tallied even when the URL comprises an *auto-login* feature through parameters specified in the URL query string (see below).

The main interest of collecting accesses is to know when a user has accessed the instance of the toolbox. As already mentioned, there is the possibility to access the EAT, but without producing any data. So the observations alone may not be reliable in determining whether users have been *passively* using the EAT.

Table 6.1 illustrates the variables recorded for each access to the instance. For each variable, a description is also provided.

Table 6.1: List of variables related to accesses, with description.

Variable	Description
Participant	ID of the participant specified through the login form or the URL of the instance
Group	ID of the group specified through the login form or the URL of the instance
Study	The unique ID of the instance, given by the toolbox
Date	The UTC date and time of the server
Query String	A string of key-value associations that are appended at the end of an URL. They can be useful to record information about participants given outside the toolbox

6.5.2 Data About Observations

Observations consists in the *unit* that is encoded every time a user send data across the expressing-displaying area of the tool. In most cases, this will correspond to an *emotional experience*. As mentioned above, though, the flexibility of the toolbox may confer to the *unit* different meaning, so the more neutral term of *observation* is adopted in the interface.

Given the self-report nature of the tool, an observation must be explicitly sent by users. In the interaction sequence, thus, the action must be terminated on a voluntary click on a button, be it a suggested lexicalized emotion, the *Other…* form, or an *alternative* button. The toolbox does not record partial or ongoing manipulations on the interface, but only the final result.

The number of variables recorded for each observation at the time of writing is around 30, and this number may increase over time as more meaningful details or automated processing of collected data can be added. Some variables also depends on the specific structure of the underlying affective space. In this section, only the most relevant variables are described. The presentation will also be divided in Cartesian plane and circumplex. More details will be available in the technical documentation referenced in the Code and Data section at the end of the Introduction.

For data relative to Cartesian plane's affective spaces (uni-, bi- or multi-dimensional), each observation is characterized by the variables illustrated in Table 6.2. A description provide some additional information.

Table 6.2: List of variables related to observations when the affective space is an N-dimensional Cartesian plane

Variable	Description
Participant	ID of the participant specified through the login form or the URL of the instance
Group	ID of the group specified through the login form or the URL of the instance
Study	The unique ID of the instance, given by the toolbox
Date	The absolute UTC date and time of the server
Elapsed time	The relative time elapsed from the start of a session (see the section about accesses above)
Dimension(s)	One variable labelled with the name of each dimension of the affective space. The value corresponds to the rating effectated on the corresponding slider on the interface
Last dimension	The last dimension whose slider has been modified before sending the observation
Click type	This may take one of three values: button, other, or none. The first value corresponds to the click on a suggested button, the second to the use of the Other field, and the last the alternative/none button
Label	The exact term(s) corresponding to the subjective feeling. Depending on the click type, that can be a lexicalized emotion belonging to the affective space, a term provided by the user, or the label of the alternative/none button
Listed	A true or false variable that defines whether the proposed label is included in the pre-defined list of lexicalized emotion of the affective space
Label sanitized	The given label, stripped of special character (useful especially in languages such as French or Italian)
Closest label	Using a string-comparison algorithm, the pre-defined lexicalized emotion in the affective space which is closest to the given label

Distance closest	A quantification of the distance between the exact label and the closest, where 0 means exact match
Expected dimension(s)	If a match of the exact label is found in the pre-defined list of the affective space, the expected value on each dimension is retrieved from the reference frame
Vector distance	If a match of the exact label is found in the pre-defined list of the afftive space, the computed Euclidean distance between the rated dimensions and the expected dimensions is calculated
Online users	A list of IDs of the connected users. The count of online users is also provided in another variable.

The sense and use of most variables is straightforward: they are extension or refinements of the core dimensional and discrete emotions approaches to emotion self-report. The most relevant variables from a theoretical perspectives are the list of online users and the vector distance.

The list of online users is meant to enhance measures related to the social presence (Jézégou, 2010; Kreijns et al., 2013; Lowenthal & Snelson, 2017), illustrated in Section 1.3.3. For instance, in a longitudinal use of the EAT, it may be interesting to investigate whether the learners are more prone to share their emotions when there are other colleagues connected at the same time. More fine-grained analysis could also reveal whether specific learners have an impact on the overall emotional experience (e.g. when *P123456* is active, there is an overall increase in the x dimension).

The vector distance quantifies $\Delta(E, P)$, where E is represented by the *Dimenson(s)* variables, whereas P is represented by the *Expected dimension(s)* variables. As stated in the description, this variable can only be computed if the exact label of the subjective feeling is detected as being part of the pre-defined list. Future version of the toolbox may in this sense use the closest label conditional on a predefined threshold of the string-comparison algorithm, but this is not the case at the time of writing. This kind of computation can in any case be performed afterwards, so they are meant to be shortcuts and helpers, not indispensable features.

For a bi-dimensional circumplex, the variables are mostly the same. The radial distance is replaced by similar measures quantifying $\Delta(E, P)$. These variables are described in Table 6.3. All the other variables common with the Cartesian plane are omitted.

Table 6.3: List of variables related to observations when the affective space is a bi-dimensional circumplex. Common variables with Cartesian plane's spaces are omitted.

Variable	Description
Angle	The computed angle, using the arctangent, retrieved from the two-dimensional values

Observed quadrant	What quadrant of the circumplex corresponds to the two-dimensional values
Expected angle	If a match of the exact label is found in the pre-defined list of the affective space, the expected angle of the lexicalized emotion in the circumplex
Expected quadrant	If a match of the exact label is found in the pre-defined list of the affective space, the expected quadrant the lexicalized emotion belongs to in the circumplex
Radial distance	The absolute and shortest (i.e. $< 180°$) distance between the observed and expected angle, if a match is found

The measures specific for the circumplexes are simply adaptations to the use of angles rather then the Euclidean distance. For instance, the radial distance must be congruent with the continuity of the circle, meaning that there are the clockwise and anti-clockwise routes between any two points on the circumplex. That is, if a lexicalized emotion in the list has a theoretically-driven position of 15°, but is evaluated to the equivalent of 300°, the shortest radial difference is 75° (anti-clockwise) and not 345° (clockwise).

The overall purpose of the observed vs. expected measures, both for Cartesian plane's and circumplexes, is the possibility to compare the theoretically-driven (expected) affective space, with the one empirically rated (observed) by users. For instance, it would be possible to compute the mean of the ratings for any listed lexicalized emotion as an empirically determined position on the affective space. This mechanism is used Chapter **??**.

To sum up, the measures available for the observations recorded through the toolbox are meant to facilitate meaning-making about the overall emotional experience of participants or learners. The measures propose some basic pre-processing, meant to improve the quality of the data or facilitate some possible analysis.

6.6 Summary

This chapter presented a toolbox built around the computational model at the core of the Dynamic Emotion Wheel. The toolbox is meant to fulfill, as best as possible, the multi-purpose vocation of the EAT. Researchers and practitioners can configure many aspects of the interface, which may be related to specific theoretical, empirical, or pedagogical interests. At the same time, the toolbox is also meant to foster open science best practices in transparency, reproducibilty and replicability of results, as well as sharing of empirical material and data.

The toolbox is used in the empirical contributions presented in the following Part III of the thesis. It is worth reminding in this regard, though, that what has been presented in this chapter is the latest cycle of the iterative development of the software. Some features have been added as a direct results of what has been observed in the empirical contributions (e.g. "*I wished I had this feature. . .* " or"*The toolbox must fix this problem*"). At the same time, not all features were available at the time of planning the studies and while collecting the data.

Part III

Empirical Contribution

Chapter 7

Does a Different Use of and Access to Emotional Information Change the Concrete Use of the Emotion Awareness Tool?

This chapter describes the first empirical contribution of the thesis, which investigates to what extent a different use of, and access to emotional information can determine the concrete use of an EAT in synchronous and collaborative computer-mediated settings. The presentation of the experiment follows the traditional structure of empirical contributions (Sollaci & Pereira, 2004) and therefore the outline of the chapter is not presented. With respect to a traditional contribution, though, the theoretical background consists only in reminders of theories or empirical contributions already depicted in Part I, which are directly integrated into the rationale of the study presented at the beginning of the chapter.

7.1 Study Rationale

In the few experimental contributions that have investigated emotional awareness in synchronous computer-mediated collaboration so far (*e.g.*, Avry, Molinari, et al., 2020; Eligio et al., 2012; Gaëlle Molinari, Chanel, et al., 2013), the experimental design consisted in comparing a control group, who did not dispose of emotional awareness, with a *treatment* group, who disposed of emotional awareness in *full strength*. That is, either the person had no emotional awareness at all, or she had access to the full *abstract model* depicted in Section 2.5: intra-individual and inter-individual awareness, as well as the comparison between the two. The *treatment* setting is therefore consistent with the mutual-modeling perspective (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; M. Sangin, 2009), according to which the symmetry of the information available to learners is instrumental to build and update a holistic representation of the partner, upon which the collaborative effort can strive.

As the abstract model has shown, though, emotional awareness can be broken down to a series of passages that, even though tightly related, may have different reasons for, and consequences on the use of an EAT. At the very least, it is possible to identify the intra-personal and the inter-personal *paths*, which are underpinned by two different assumptions about the instrumentality of emotional awareness (see Sections 2.2 and 2.3). In this regard, for instance, Gaëlle Molinari, Chanel, et al. (2013) reckon that "one main limitation of [their *control vs. treatment*] study is the difficulty in disentangling the effect of reflecting upon one's own emotions from the effect of sharing one's emotions with the partner" (Gaëlle Molinari, Chanel, et al., 2013, p. 342). As a consequence, it may be worth investigating the use of different configurations of an EAT, which vary, for instance, according to (1) the use of the emotional information that is produced, and (2) the access to the emotional information that may be perused through the tool. As suggested by Buder (2011), varying the characteristics within an awareness tool may complement the assessment of its specific contribution alongside the all-or-nothing approach of the *control vs. treatment* design.

One way to guide the configuration of different versions of an EAT is to use the abstract model of the functions of emotional awareness and imagine it is a sort of *pipeline* in which the *flow* of emotional awareness circulates. Each passage from one main concept – the learning task, the intra-indiviudal or inter-individual emotion, and the meaning-making extrapolated from emotional information – can be seen as a pipe, whose flow can be blocked, decreased, or increased according to the configuration of the EAT. Figure 7.1 shows this concept graphically, with three different configurations of the *pipeline*, identified respectively as Self-Centered, Partner-Oriented, and Mutual-Modeling.

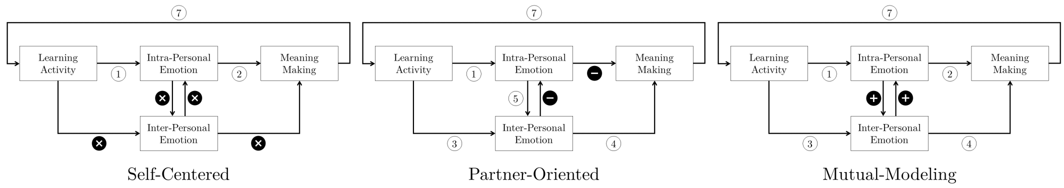


Figure 7.1: Comparison of three different *pipelines* in the abstract model of emotional awareness: the Self-Centered, the Partner-Oriented, and the Mutual-Modeling *paths*. The *flow* in a pipe can be blocked (x), decreased (–), or increased (+), depending on the configuration of the the EAT's interface.

In the Self-Centered *path*, all the passages related to the inter-personal emotion edge are blocked, whether incoming or outgoing. Conceptually, this means that the learner is provided only with emotional self-awareness, since the emotion expressed by each participant to the collaborative task remain *private*: the emotional information is not disclosed to others, and therefore each learner peruse only the emotional information she has produced.

In the Partner-Oriented *path*, all passages remain open, but the *flow* is reduced in passages 2 (from intra-personal emotion to meaning-making) and 6 (from inter-personal to intra-personal emotion). Conceptually, this means that the emotion expressed by the learner are disclosed to the partner, but they do not persist on the interface of the learner herself as a means to foster intra-individual meaning-making. The only available emotional information to peruse for the learner consists in the emotions of the partner. This entails that direct and persistent comparison between the partner's emotions and that of the learner is considerably reduced: it would depend only on the learner memory of her emotional experience over time.

Finally, in the Mutual-Modeling *path*, all the passages remain open, and the *flow* between the intra-personal and inter-personal emotion edges is *increased* by the direct and persistent comparison between the emotional experience of the learner herself and that of the partner. Conceptually, this entails a total symmetry between the emotional information available to both partners, which is consistent with the mutual-modeling activity through which the holistic representation of the partners is built and updated (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; M. Sangin, 2009).

The analogy of the pipeline is also relevant to address two important issues in the study of emotional awareness when the inter-personal perspective is implicated. First, the overall *flow* is influenced by how much information is produced by each learner. That is, the more emotions are shared by the learner, the more the partner will dispose of emotional information to peruse. At the same time, the more emotions learner share respectively, the more the symmetry will be quantitatively and qualitatively achieved. It could in fact be the case that even if learners dispose of the Mutual-Modeling interface, one shares a lot of emotions and the other none at all. Second, as indicated by Janssen & Bodemer (2013), the content and the relational space of a computer-mediated collaboration overlap (see Figure 1.1 in Section 1.3.4). In other words, it is possible to infer emotional information that comes from outside the EAT, and also produce emotional information using other sources than the EAT. For instance, one can infer that a colleague who is not writing in a shared document, but who follows with the cursor the new lines added, is probably uneasy about how to contribute. Thus, if one is interested in the investigation of the specific contribution of an EAT, other sources of emotion-related information shall be accounted for or controlled in some way.

In order to address both issues, the contribution will therefore adopt a simulated collaborative setting as the one used in the usability test of the Dynamic Emotion Wheel illustrated in Section 4.3.4. This will expose all participants to the same *content* of the collaboration, whereas the *relational* information will be different with respect to the emotional information available (Janssen & Bodemer, 2013).

7.2 Research Question and Hypothesis

The main aim of the experiment is therefore to manipulate the abstract model of the functions of emotional awareness and determine what does it happen – holding all other sources of potential variation in the emotional information constant – if the *flow* of emotional information is blocked, reduced, or increased in specific parts of the model. Depending on the use that would be made of the emotions expressed, as well as to what kind of emotional information learners have access to, there are different reasons to *encode* and *decode* emotional information into and from the EAT. This is all the more relevant with a self-reporting and moment-to-moment use of an EAT. As stated by the Component Process Model (Scherer, 2005, 2009b, 2019a) for the intra-individual level, and by the Emotion As Social Information model (Van Kleef, 2009, 2010, 2018) for the inter-individual, producing and consuming emotional information requires a cognitive effort when the process happens at the conceptual and inferential levels. This effort must be evaluated according to the trade-off between the cost of producing or perusing the emotional information on the one hand, and the benefit of disposing or making the others dispose of emotional awareness as instrumental information to the task at hand (Buder, 2011; Dillenbourg et al., 2016; Pashler, 1994).

In this regard, one way to look at the three different *flows* of emotional awareness is to consider that, starting from the Self-Centered *flow* and ending with the *Mutual-Modeling* flow, there is on every step an increment on how socially oriented the interface is. Figure 7.2 shows this mechanism graphically by mobilizing concepts illustrated in Chapter 2.

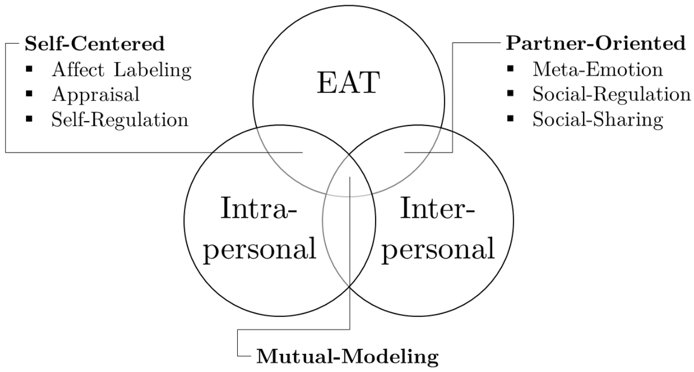


Figure 7.2: Theoretical concepts mobilized by versions of the EAT differing in the use of, and access to emotional information. Concepts are listed in alphabetic order.

In the Self-Centered condition, learners will be oriented towards their own emotional experience (Lavoué et al., 2020; Gaëlle Molinari et al., 2016). This translates, for instance, in an accurate appraisal of the situation (Scherer, 2021a), which is nevertheless not influenced by the need to reflect on a strategic signal to others or normative pressure (Hareli et al., 2013; Scherer, 2007). Another important process in a Self-Centered orientation is represented by emotion self-regulation (J. J. Gross, 2014; J. J. Gross, 2015): expressing and perusing one's own emotions through the EAT can enhance self-reflection on the emotional experience, and therefore strategies to maintain functional emotions and modify dysfunctional ones. In this regard, the use of an EAT linking appraisal dimensions and subjective feelings can enhance a form of regulation through *affect labeling* (Lieberman et al., 2007; Lieberman, 2019b; Torre & Lieberman, 2018), according to which the process of *naming* an emotion with a term entails an implicit regulation of that same emotion.

In the Partner-Oriented *flow*, learners will be concerned by the disclosure of their emotion to the partner and by taking the emotions of the partner into account. All the implications of the Self-Centered condition are maintained, but also implemented by the fact that emotional awareness emerges as a bi-directional communicative act. It therefore entails all the considerations regarding the social-sharing of emotion (A. H. Fischer & Manstead, 2016; Parkinson, 2008; Rimé, 2009; Van Kleef, 2018). In addition, if a form of emotion regulation is implied, that could also be from an inter-personal perspective (Netzer et al., 2015; Reeck et al., 2016; Zaki & Williams, 2013). Finally, the emotion expressed by the partner can become a trigger for an emotion in the learner, that is, an inter-personal meta-emotion (Miceli & Castelfranchi, 2019; E. Norman & Furnes, 2016).

Finally, in the Mutual-Modeling *flow*, all the previous concepts coalesce, with the addition of direct and persistent comparison of the evolution of the emotional experience in the learner and in the partner (Avry, Molinari, et al., 2020; Eligio et al., 2012; Gaëlle Molinari, Chanel, et al., 2013). This visuo-spatial comparison enhance the symmetry of the partner modeling (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; M. Sangin, 2009), which is considered a pivotal mechanism for inter-subjective meaning making in collaborative settings (Janssen & Bodemer, 2013; Järvelä et al., 2015; Kirschner et al., 2015; Phielix, Prins, Kirschner, Erkens, & Jaspers, 2011).

The main research question in this experiment consists thus in assessing whether a different use of and access to emotional information changes the concrete use of an emotion awareness tool. The use of the tool is determined by the two fundamental functions of awareness tools: the expressing-displaying function (producing awareness) and the perceiving-monitoring function (perusing awareness). For each function, specific hypotheses and causal mechanisms are depicted hereafter.

7.2.1 Use in Expressing Emotions

For expressing-displaying emotions, the three interfaces provide the learner with different reasons to express how she feels, as well as different affective triggers that may elicit emotional episodes (see also the hypothesis about the use in perceiving emotions below). More precisely:

- With the *Self-Centered* interface, the learner knows she is the only recipient of the emotions she expresses. Therefore, if she decides to use the EAT to express an emotion, she probably does it out of self-interest, possibly linked to self-regulation as stated by the *intra-personal* perspective. In the meantime, the EAT does not provide any additional information about the partner's emotions that may serve as a trigger for emotional episodes in the learner herself.
- With the *Partner-Oriented* interface, the learner knows she will not have access to her own emotions once she has expressed them, but that these are conveyed to the partner. Therefore, if she decides to use the EAT to express an emotion, one can assume that she does it from an *inter-personal* perspective (even if the possibility that she does it exclusively in a *Self-Centered* perspective cannot be excluded). In the meantime, the learner can also access the partner's emotions, which may represent additional triggers for meta-emotional episodes (*e.g.*, *Jane expresses guilt because she thinks Paul has just expressed anger as a result of something she has done*).
- With the *Mutual-Modeling* interface, the learner knows the emotions she expresses are available both to her and the partner. Therefore, if she decides to use the EAT to express an emotion, she does in a *Self-Centered*, *Partner-Oriented*, or a combined perspective. In the meantime, the learner also disposes of direct and persistent comparison between her own emotions and that of the partner, which may also represent an additional trigger for emotional episodes compared to the *Partner-Oriented* inter-

face (*e.g.*, *Jane expresses relief because she saw from the interface that in the last few minutes both she and Paul were often confused*).

Hypothesis (*H1*) is therefore stated in the following terms: there will be an overall difference in the use of the EAT for expressing-displaying emotions depending on the interface the learner has at disposal. More specifically, in comparing the interfaces, a greater use of the expressing-displaying function of the EAT in the *Partner-Oriented* and *Mutual-Modeling* interfaces compared to the *Self-Centered* interface would corroborate an *inter-personal* interest in expressing emotions. Furthermore, a greater number of emotions expressed in the *Mutual-Modeling* interface compared to the *Partner-Oriented* condition would suggest that the possibility of direct and persistent comparison between one's own and the partner's emotions results in a *surplus* of expression-displaying of emotions.

7.2.2 Use in Perceiving Emotions

With respect to perceiving emotions, the three interfaces differ in the quality and quantity of the emotional information available on screen. The three interfaces will provide the learner with different reasons to seek and process the emotions expressed during the collaboration:

- With the *Self-Centered* interface, the learner has access only to the emotions she has expressed over time during the collaboration. This may be interpreted as a *control* condition: what is the interest of having emotional information that the learner is already supposed to know? Seeking and processing the learner's own emotions may be explained by the interest of reflecting on the evolution of her own affective states during the task.
- With the *Partner-Oriented* interface, the learner has access only to the emotions expressed by the partner, that is, information she does not already know. Seeking and processing the partner's emotions may be explained by the interest in knowing how the other is feeling and/or the evolution of the affective states of the partner during the collaboration.
- With the *Mutual-Modeling* interface, the learner has access both to her own and the partner's emotions. This condition inserts an additional interest to the previous ones: the possibility of direct and persistent comparison between the learner's own emotions and that of the partner.

Hypothesis (*H2*) is therefore posited as follows: there will be an overall difference in the use of the EAT for seeking and processing the expressed emotions depending on the interface the learner has at disposal. More specifically, the *Partner-Oriented* and *Mutual-Modeling* interfaces will elicit a greater use in perceiving-monitoring emotions compared with the *Self-Centered* interface. Furthermore, greater information seeking and processing in the *Mutual-Modeling* interface compared to the *Partner-Oriented* interface would suggest an accrued interest due to direct and persistent comparison.

7.3 Methods

Following Simmons, Nelson, & Simonsohn (2012) suggestion to increase transparency in experimental contributions, I report how I determined the sample size, all data exclusions (if any), all manipulations, and all measures in the study.

7.3.1 Participants and Design

N = 48 participants (29 women, 19 men), aged 18 to 55 (M_{age} = 37.3, SD_{age} = 10.01), voluntarily participated to the study. The sample size was determined by time constraints, since data had to be collected in 15 days in co-located conditions. 23 participants were university students from different faculties, both at undergraduate and graduate levels. 25 participants were professionals working for a company adopting distance learning practices. No remuneration was provided for taking part in the study. Participants were randomly assigned to one of the three conditions/interfaces (*Self-Centered*, *Partner-Oriented*, or *Mutual-Modeling*) in order to produce a balanced design with 16 participants per condition.

7.3.2 Material

Overall Interface of the Task

The experimental material comprises different components; therefore an overview is provided before specifying the various details. Figure 7.3 shows the disposition of the screen during the experimental task. It comprises the EAT on the left-side of the screen, outlined in blue, and the simulated, joint-problem solving task on the right. The image indicates what part of the interface was simulated for the *Mutual-Modeling* condition. The interfaces of the other conditions are illustrated below. Some parts of the interface have been translated in English for the current contribution. In the experiment, though, the french language was used consistently for every condition.

Problem-Solving Task

The joint problem-solving task comprised four enigmas taken from a game. The first three enigmas had a clear response that could be inferred, whereas the last one was more of a non-nonsensical type. The same enigmas have been used in the usability test of the Dynamic Emotion Wheel (Fritz, 2015), where they elicited different emotions both in number and kind in a population similar to that of the current sample (see Section 4.3.5). Each enigma was subdivided in three phases:

- 40 seconds during which the text of the enigma was showed on the interface. At this stage, participants could only express their emotions, but could not write their reasoning or the reply;
- 3 minutes and 20 seconds during which participants could write their reasoning and reply to the enigma, as well as see the *playback* reasoning (but not yet the answer) of the simulated partner;

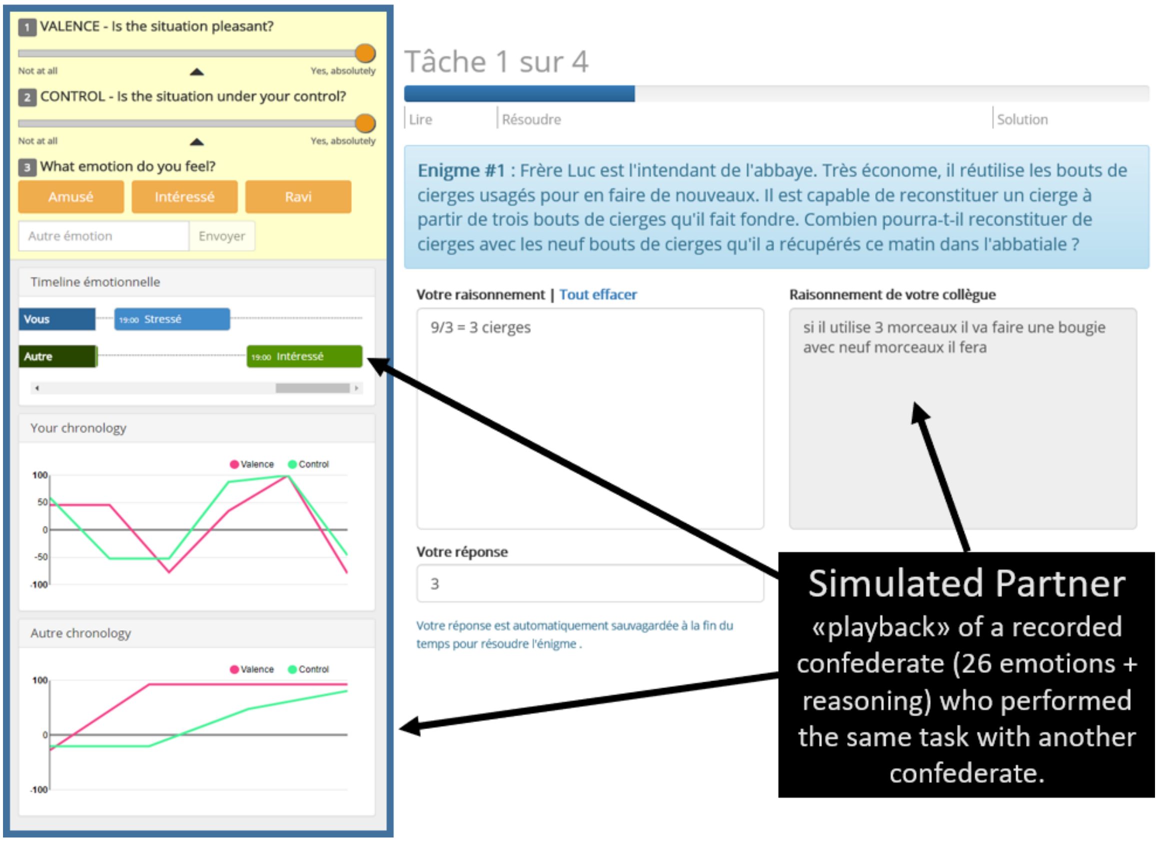


Figure 7.3: Overview of the interface that presents the various part of the material adopted for the CSCL task.

- 1 minute in which the given answers from the participant and the partner were displayed on screen with the expected solution to the enigma. At this stage, once again, the reasoning and reply fields were not available on screen.

Configuration of the Emotion Awareness Tool

The toolbox was adopted to manipulate the interface in order to obtain the three different conditions. The expressing-displaying function was common to all conditions, which in turn differed in the perceiving-monitoring part of the interface.

For expressing an emotion, the underlying affective space was represented by EATMINT circumplex depicted at length in Section 4.3.3, which is determined by the *Valence* and *Control/Power* appraisals (Scherer, 2005; Scherer, Shuman, et al., 2013). Valence was prompted with the question *Is the situation pleasant?* *Control/Power* was evaluated with the question *Is the situation under your control?* Both evaluations were determined with the extreme negative pole *Not at all*, and extreme positive pole *Yes, absolutely*.

For the perceiving-monitoring function of the EAT, each condition differed in the following ways (depicted in Figure 7.4):

- *Self-Centered*: the interface comprises an emotion time-line, then a line chart that depicts the evolution of the appraisal dimensions over time, and finally a tag cloud where the size of each subjective feelings is proportional to the frequency with which it has been expressed. The information provided is based only on the emotions expressed by the participant herself.

- *Partner-Oriented*: the interface mirrors that of the *Self-Centered* condition, for the information provided is based only on the emotions expressed by the simulated partner (see below).
- *Mutual-Modeling*: the interface comprises an emotion time-line, but with both the participant and the simulated partner's subjective feelings organized in two different rows. Two line charts complete the interface, one with the appraisal dimensions of the participant, and the other of the partner.

The *Self-Centered* and *Partner-Oriented* conditions present a tag cloud at the end of the interface in order to balance the surface of the EAT that contains information. In this way, the EAT occupies more or less the same amount of the screen.

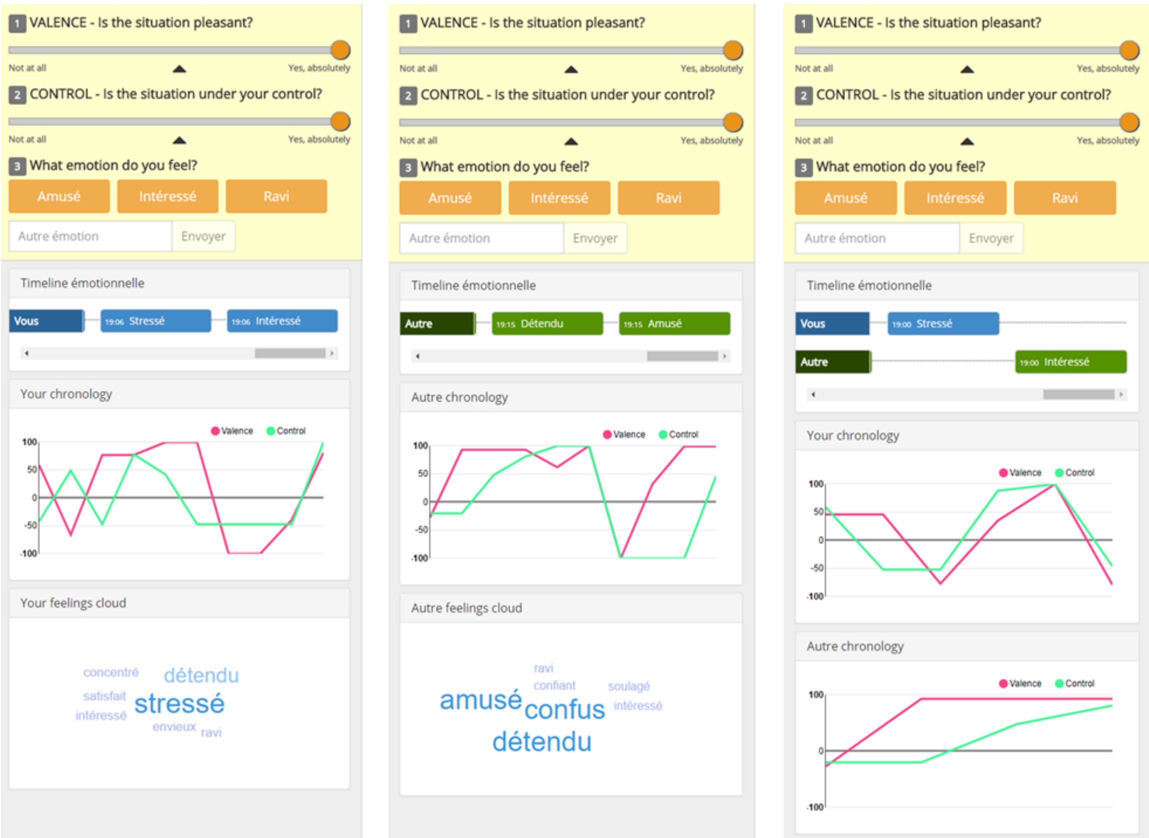


Figure 7.4: The three different interfaces used in the experiment. From left to right: the *Self-Centered*, the *Partner-Oriented*, and the *Mutual-Modeling* versions.

Simulated partner

The *playback* manipulations displayed on the interface were recorded in the same pilot test as for the usability assessment of the DEW already mentioned above. 4 confederates (2 men and 2 women) performed the same joint problem-solving task, but in a synchronous situation. The *playback* is thus comprised by: (1) all the emotional episodes expressed, represented by the evaluation on the two appraisal dimensions and the related subjective feeling; and (2) what the confederate have typed, at the very same moment, into the reasoning field, as well as the answer to each of the 4 enigmas. In this regard, confederates were explicitly asked not to communicate directly through the text fields, but limit their typing to the reasoning for solving the problem. One of the *playback* was then randomly chosen for the task and *injected* into the experimental task interface. The simulated partner expresses 26 emotions and finds the solution to 2 out of 4 enigmas. The full list of emotions – comprising the time of expression, the associated *Valence* and *Control/Power* appraisals, and the subjective feeling – are depicted in Table 7.1.

Table 7.1: List of the emotions of the simulated partner. Time is expressed in seconds.

Time	Valence	Control	Feeling (FR)	Feeling (EN)
29	-27	-20	Confus	Confused
100	93	-20	Détendu	Relaxed
118	93	48	Amusé	Amused
153	93	81	Intéressé	Interested
236	62	100	Confiant	Confident
253	100	100	Ravi	Delighted
321	-100	-100	Confus	Confused
383	32	-100	Soulagé	Relieved
414	99	-100	Détendu	Relaxed
430	99	46	Amusé	Amused
520	100	100	Ravi	Delighted
554	100	100	Ravi	Delighted
572	100	100	Satisfait	Satisfied
610	-100	-100	Frustré	Frustrated
683	-100	-100	Lassé	Bored
746	-100	28	Insatisfait	Disappointed
815	-100	-40	Frustré	Frustrated
863	-100	-100	Insatisfait	Disappointed
877	-75	-72	Lassé	Bored
932	-100	-100	Anxieux	Anxious
951	-100	-100	Confus	Confused
1013	-100	-100	Frustré	Frustrated
1031	-100	-46	Amusé	Amused
1159	-100	-61	Insatisfait	Disappointed
1168	-100	-100	Frustré	Frustrated
1173	-100	-100	Confus	Confused

Eye-tracking

A Tobii T120 eye-tracker with Tobii Studio Pro v3.4.8 software (Tobii AB, 2015) was used for eye-tracking measures. Areas Of Interest (AOI) were disposed on the EAT as a whole (left side of the screen) and the task (right side of the screen). The AOI of the EAT was further divided in the expressing-displaying upper zone, and the perceiving-monitoring lower zone.

7.3.3 Procedure

Participants were given a specific time to come to the test, which was performed at Geneva University, and were reminded of the importance to be on time since another participant was performing the test in the meantime. The experimenter welcomed the participant in the room with the eye-tracking equipment. Once installed, the experimenter proceeded to explain the outline of the study:

- Introduction and explication (10 minutes)
- Warm-up session with the DEW and instructions for the task (5 minutes)
- Collaborative task (20 minutes)
- Debriefing (10 minutes)

Introduction and explication

The general aim of the study was explained. The experimenter reassured participants about the fact that the data would be anonymous, and that they could stop the experiment at any time without any reason. A first consent form was then signed if the participant agreed to take part in the study.

At this point, the experimenter explained how the collaborative task would take place. She first illustrated a demo about the functioning of the DEW. Since in a previous study (Fritz, 2015), whose aim was to observe the spontaneous use of the tool, participants were confused about the dimension of *Control/Power*, in this experiment the experimenter proposed a more thorough explanation of what the two sliders of the DEW stand for. The explication also aimed at reducing the risk that participants will move the cursors for the *Valence* and *Control/Power* dimensions until they found the *right* subjective feeling. Next, the experimenter showed the perceiving-monitoring part of the EAT, which was explained according to the experimental condition the participant was attributed to. In this way, participants were informed about both what the emotional information they provided would be used for (*Self-Centered* vs *Partner-oriented/Mutual-modeling*), and to what kind of emotional information they would have access to (*Self-Centered* vs *Partner-Oriented* vs *Mutual-Modeling*).

Finally, the experimenter explained the right-hand side of the screen, which implemented the joint problem-solving task. Participants were informed about the three parts (reading, solving and solution) that composed each of the 4 enigmas to solve. They were also prompted to write their reasoning to solve the problem in the appropriate field, but to avoid direct communication with the partner. Since in the usability test of the DEW participants spontaneously expressed emotions and participated to the task, a system of point to foster collaboration originally adopted was dropped.

Warm-up session with the DEW

Participants were placed in front of the screen used for the test and could practice with a simplified version of the interface for the task. Participants could familiarize with the expression of emotions through the DEW, and random emotions were also injected in the interface at short intervals to emulate the emotion of the partner. The side of the screen devoted to the task was filled in with generic texts explaining what participants will see in the actual task (*e.g.*, *here will appear the text of the enigma*, *here you must write your reasoning*, . . .)

Experimental task

Once the participant was ready for the test, the experimenter simulated to check-in with another confederate to simulate that the other participant was ready to start the experiment as well. Then the experimenter proceeded to calibrate the eye-tracker equipment. After being reminded about the general functioning of the eye-tracker and the importance of not moving during the task, the participant would then proceed with the task. At first, she had to fill the *log-in form* of the front-end interface of the toolbox, providing a random ID and an identifier for the pair. The system simulated a synchronization latency time, for the other participant to reach the same stage. Once the task started, the participant had access to the overall interface depicted above, including the *playback* of all the manipulations made by a confederate.

Post-test debriefing

After the task, participants were asked to fill in a survey whose data will not be used in the present contribution, but have been analyzed in another contribution (Perrier, 2017). At the end of the study, participants were informed about the manipulation of the simulated partner and the experimental reasons behind it. A second consent form was therefore submitted to participants, for them to confirm they understood the reason of the manipulation, and that they accepted the use of the data for scientific purposes.

7.3.4 A Priori Exclusion Criteria

Exclusion criteria determined beforehand concerned only technical issues that could jeopardize the task, especially with respect to the simulated partner. Any interruption of the task or technical failure would make the trial not recoverable. Exclusion caused for low quality of eye-tracking measures, due for instance to the participant moving too much, were also foreseen, but not yet quantified due the lack of a precise benchmark.

7.3.5 Data analysis

For hypothesis *H1*, concerning differences in expression of emotions, a omnibus one-way ANOVA with pairwise comparison between all conditions was planned beforehand. The number of emotions expressed through the EAT represents the dependent variable, and the interface of the EAT the independent variable with the three conditions (*Self-Centered*, *Partner-Oriented*, or *Mutual-Modeling*).

For hypothesis *H2*, concerning differences in perception of emotions, two indicators retrieved by eye-tracking measures (Blascheck et al., 2017; Poole & Ball, 2005) are used as dependent variables. First, the number of times participants sought information by orienting their gaze inside the perceiving-monitoring zone of the interface, which is usually interpreted as an indicator that the person intentionally seeks for information she may find useful or that she got lost and needs reorientation. Second, the total time (in seconds) that participants spent looking at the perceiving-monitoring zone of the interface. Such indicator is usually interpreted as a proxy for information processing and could account for interest (*i.e.*, people look at it longer because it is interesting) or complexity (*i.e.*, people look at it longer because they need more time to understand what it means). Given the relative simplicity of the information provided – even though people do not like graphs (Carpenter & Shah, 1998; Pinker, 1990) – and the use of a fixed interface of the EAT, both measures are used as indicators of interest, and therefore the greater the indicator, the greater use of the EAT is inferred. For both dependent variable, a omnibus one-way ANOVA with pairwise comparison between the three conditions of the independent variable (*Self-Centered*, *Partner-Oriented*, or *Mutual-Modeling*) were planned beforehand. A family-wise correction to account for inflation in Type I error has been planned for each pairwise comparison.

All analysis are conducted using the statistical software R, version 4.2.0. Analysis of variance use the Afex package version 1.1.1 (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2020).

7.4 Results For Planned Analyses

In this section, only the results of the planned analysis are provided. A subsequent section describes post-hoc, corollary analysis that were not planned beforehand.

7.4.1 Post-Hoc Exclusion

Results will be based on $N = 35$ participants. 10 participants were excluded due to technical issues during the task or low quality of eye-tracking measures. One participant was excluded for statistical reasons: the participant expressed 62 emotions during the task, against a mean of $M = 13.80$ ($SD = 5.68$), that is, more than 8 standard deviations above the mean. Such a number, not even close to any participant to the same task in Fritz (2015), suggests a non representative use of the tool. The distribution of participants after post-hoc exclusions with respect to the experimental conditions is depicted in Table 7.2.

Table 7.2: Number of participants retained for each experimental condition ($N = 35$).

Condition	N
Self	12
Partner	9
Mutual	14

The resulting unbalanced design and overall small N, particularly low in the *Partner-Oriented* condition, decrease the power of an already ambitious planned test and make it more exposed to violation of assumptions (Hoekstra, Kiers, & Johnson, 2012). The interpretation of results from an hypothesis-testing perspective is nevertheless maintained, but provided from an exploratory perspective, such as a reference for future studies, rather than a confirmatory standpoint (Fiedler, 2004; Lakens & Etz, 2017; Scheel et al., 2020b).

7.4.2 Differences in Expressing Emotions

Participants expressed a total of 483 emotions, which corresponds to a mean close to 14 emotions per participants ($M = 13.80$, 95% CI $[11.85, 15.75]$, $SD = 5.68$). As a reference, in the usability test (see Section 4.3.5) the average was of $M = 17.14$, with a similar standard deviation of $SD = 5.05$. Participants in the *Self-Centered* condition expressed on average $M = 12.50$ ($SD = 5.30$) emotions; $M = 12.67$ ($SD = 6.00$) in the *Partner-Oriented* condition; and $M = 15.64$ ($SD = 5.69$) in the *Mutual-Modeling* condition (see Figure 7.5).

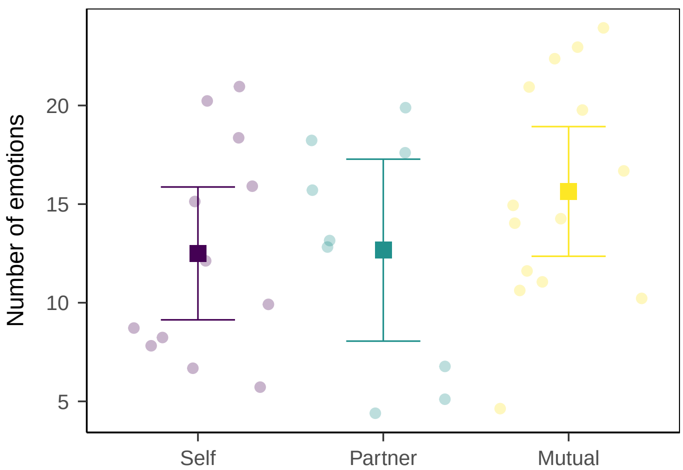


Figure 7.5: Number of emotions expressed by experimental condition. Bars represent 95% confidence intervals.

An overall effect of the interface on the number of emotion expressed could not be detected neither in the omnibus ANOVA ($F(2, 32) = 1.25$, $p = .301$, $\hat{\eta}^2_G = .072$, 90% CI $[.000, .221]$), nor in the pairwise comparisons, which are illustrated in Table 7.3. Hypothesis *H1* is therefore not supported by the data at hand. In any case, the check of the assumptions of the ANOVA model revealed a problem with the normal distribution of residuals (Hoekstra et al., 2012). Results would have been thus withdrawn even if a difference could be detected.

Table 7.3: Pairwise comparisons of the three interfaces with respect to the number of emotions expressed during the task (p-values are adjusted with the Tukey method).

Comp.	Estimation	SE	df	t	p	ΔM
Self - Partner	-0.17 [-6.28, 5.95]	2.49	32	-0.07	.998	-0.03 [-0.93, 0.87]
Self - Mutual	-3.14 [-8.60, 2.31]	2.22	32	-1.42	.345	-0.56 [-1.37, 0.26]
Partner - Mutual	-2.98 [-8.90, 2.95]	2.41	32	-1.23	.442	-0.53 [-1.41, 0.35]

7.4.3 Differences in Perceiving Emotions

Before analyzing the results for the specific hypotheses implicating eye-tracking measures, it is worth looking at overall measures about participants' gaze during the task. For instance, the information inside the AOI with the problem solving task was processed with an average of M = NA (SD = NA) seconds, which is roughly aligned with the $M = 650.61$ ($SD = 116.53$) observed in Fritz (2015), see again Section 4.3.5. All retained participants spent at least 7 minutes with their gaze intercepted inside the task area, which suggests that the task was performed as expected. The visits duration on the whole EAT have been of $M = 205.71$ ($SD = 64.01$) seconds, thus for a shorter time compared to the $M = 231.55$ ($SD = 70.86$) of the usability test. All things considering, though, the measures do not seem to deviate too much from the usability test, which may be considered as reassuring. The analysis about the perception of the emotional information is therefore presented hereafter, divided in emotional information seeking and emotional information processing.

Seeking Emotional Information

Participants' gaze entered the perceiving-monitoring zone of the interface on average $M = 67.29$, 95% CI $[55.28, 79.29]$ ($SD = 34.95$) times. In the *Self-Centered* condition, the number of visits has been $M = 40.92$ ($SD = 16.80$), whereas the count roughly doubles in the *Partner-Oriented* ($M = 75.44$, $SD = 47.35$) and the *Mutual-Modeling* ($M = 84.64$, $SD = 23.76$) conditions, for which the count was similar (see Figure 7.6).

An overall effect of the interface adopted on emotional information seeking could be detected ($F(2, 32) = 7.42$, $p = .002$, $\hat{\eta}^2_G = .317$, 90% CI $[.092, .490]$ in a one-way ANOVA). Pairwise comparisons, depicted in Table 7.4, confirm detectable differences between the *Self-Centered* vs *Partner-Oriented*, and *Self-Centered* vs *Mutual-Modeling* conditions, but not between the *Partner-Oriented* and *Mutual-Modeling* conditions. Hypothesis (*H2*) is therefore partially corroborated: the overall effect is detected, but with only two out of three comparisons between conditions. Furthermore, the confidence intervals around all parameters are wide, with the lower bound of each interval approaching zero. As stated as a preamble to the results section, thus, even if the difference reaches the threshold of statistical significance, the population effect remain uncertain.

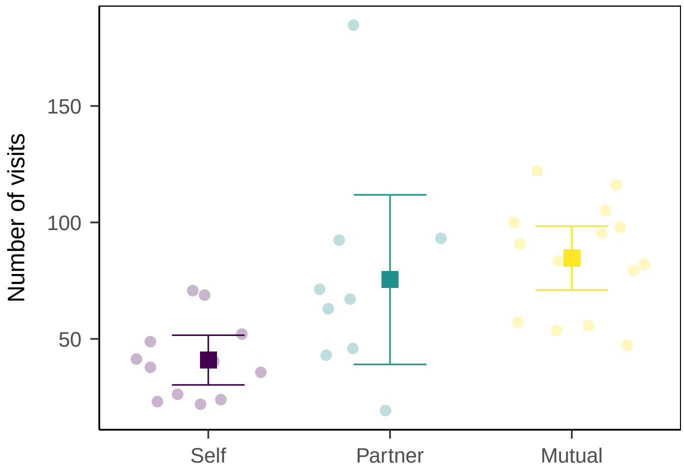


Figure 7.6: Total visits count in the perceiving-monitoring zone of the interface. Bars represent 95% confidence intervals.

Table 7.4: Pairwise comparisons of the three interfaces with respect to the number of visits at the perceiving-monitoring zone of the EAT (p-values are adjusted with the Tukey method).

Comp.	Estimation	SE	df	t	p	ΔM
Self - Partner	-34.53 [-66.80, -2.26]	13.13	32	-2.63	.034	-1.16 [-2.10, -0.21]
Self - Mutual	-43.73 [-72.51, -14.94]	11.71	32	-3.73	.002	-1.47 [-2.35, -0.58]
Partner - Mutual	-9.20 [-40.46, 22.07]	12.72	32	-0.72	.752	-0.31 [-1.18, 0.56]

Processing Emotional Information

Participants spent on average $M = 51.38$, 95% CI $[40.21, 62.56]$ ($SD = 32.54$) seconds looking at any part of the perceiving-monitoring zone of the interface, which amounts to 4.28% of the total task time. Participants in the *Self-Centered* condition spent $M = 28.32$ ($SD = 17.15$) seconds, whereas this time roughly doubles in the *Partner-Oriented* ($M = 68.53$, $SD = 39.83$) and the *Mutual-Modeling* ($M = 60.13$, $SD = 27.70$) conditions, for which time differed slightly (see Figure 7.7).

An overall effect of the experimental condition on the time spent processing emotional information could be observed ($F(2, 32) = 6.24$, $p = .005$, $\hat{\eta}^2_G = .281$, 90% CI $[.064, .457]$ in a one-way ANOVA). Pairwise comparisons, depicted in Table 7.5, confirm detectable differences between the *Self-Centered* vs *Partner-Oriented*, and *Self-Centered* vs *Mutual-Modeling* conditions, but not between the *Partner-Oriented* and *Mutual-Modeling* conditions. Hypothesis (*H2*) is therefore partially corroborated: the overall effect is detected, but with only two out of three comparisons between conditions. As for information seeking, once again the confidence interval are wide, with the lower bounds approaching zero. Results should be therefore calibrated according to the uncertainty around the population parameters.

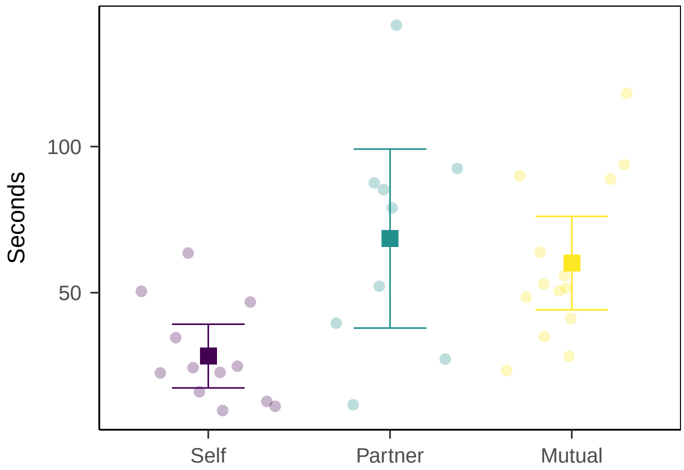


Figure 7.7: Total time (in seconds) spent looking at the perceiving-monitoring zone of the interface. Bars represent 95% confidence intervals.

Table 7.5: Pairwise comparisons of the three interfaces with respect to total time spent looking at the perceiving-monitoring zone of the EAT (*p*-values are adjusted with the Tukey method).

Comp.	Estimation	SE	df	t	p	ΔM
Self - Partner	-40.20 [-71.03, -9.37]	12.54	32	-3.20	.008	-1.41 [-2.38, -0.45]
Self - Mutual	-31.81 [-59.31, -4.31]	11.19	32	-2.84	.021	-1.12 [-1.97, -0.27]
Partner - Mutual	8.39 [-21.47, 38.26]	12.15	32	0.69	.771	0.30 [-0.58, 1.17]

7.5 Post-Hoc Corollary Analyses

In this section, I provide the results of additional analyses that have not been planned before the study. First, I extend the analysis of eye-tracking measures using transitions between Areas Of Interest as a viable measure of the use of an EAT. Second, I provide indications of the moment-to-moment use of the EAT with respect to the appraisals and subjective feeling measures collected throughout the task. Finally, I take advantage of the use of the same task as in Fritz (2015) to conduct a small internal meta-analysis (Goh et al., 2016) that can be of interest for the use of the same – or similar – task in future studies.

7.5.1 Transitions Between Areas of Interest

The eye-tracking measures used in the planned analyses of variance treated each zone of the interface as a separated element. Given the importance of dynamic, moment-to-moment phenomena in the overall thesis, it is worth investigating also transitions between the three main Areas of Interest (AOI) of the overall interface, that is (1) the expressing zone, which is common to all conditions; (2) the perceiving zone, which varies according to the experimental condition; and (3) the area dedicated to the *main* task, which is also common to all conditions. Given that transitions can go in either direction between AOI, there are 6 possible combinations of transitions: (1) Expressing to Perceiving and (2) Perceiving to Expressing; (3) Expressing to Task and (4) Task to Expressing; and finally (5) Perceiving to Task and (6) Task to Perceiving. An exploratory analysis of the transitions may reveal whether specific transitions are more frequent than others depending on the interface at disposal.

The number of transitions between AOI was computed by searching for subsequent rows in the eye-tracking logs for each of the $N = 35$ participants in which the first row had a certain AOI activated, and the following row had another AOI activated. Is it worth noting, though, that this method is sub-optimal because the experimental task included the use of mouse and keyboard. Therefore some transitions may have been lost since the gaze-path may have been interrupted by a *detour* to the peripherals. Nevertheless, it is safe to assume that participants held their mouse throughout the task, and directed their gaze into the AOI they were interested in acting upon – for instance, in order to focus the pointer into the text area – before turning the gaze away from the screen if they needed to look at the keyboard for typing. All things considering, thus, this method can be of interest at least as an exploratory method, even though it currently lacks external validity with an appropriate test. It should therefore be revisited before being deployed in a substantial analysis.

After seeing the data, one participant was excluded for having a number of transitions from Expressing to Perceiving and from Perceiving to Expressing much higher than all other participants: more than 100 against a mean of $M = 29.40$ ($SD = 13.26$) for the other participants regardless of the condition. Results are therefore based on $N = 34$ participants.

Participants made on average $M = 176.41$, 95% CI $[157.67, 195.16]$ ($SD = 53.72$) transitions between any two AOI. Figure 7.8 reports the number of transitions stratified by experimental condition between the 6 AOI organized in three rows so that each row displays the transitions between the same two AOI in each direction.

Data suggest that there are differences that may be accounted for by the type of interface the participants have access to. In particular, participants in the *Mutual-Modeling* condition seem to be more prone to make transitions between the two AOI that are more directly related to the social sharing of emotions. Transitions between the expressing-displaying and the perceiving-monitoring zone (first row in the graphic) may indicate that the possibility of direct and persistent comparison of one's own emotions with that of the partner could serve as a *social reference* before expressing one's own emotions, or *social comparison* after having expressed them. Furthermore, transitions between the perceiving-monitoring zone and the task zone (last row in the graphic) may indicate that the emotional information is taken into account as instrumental information to the task at hand. In fact, this path is the most interesting one if one consider that emotional information may be instrumental to the task at hand on a moment-to-moment basis.

Participants in the *Self-Centered* condition seem to privilege the paths between the expressing-displaying zone and the task zone, which is consistent with the fact that the perceiving-monitoring zone has only information about their own emotions. It is interesting to notice, though, that in the *Self-Centered* condition, transitions from the expressing-displaying zone to the perceiving-monitoring zone (first row, graph on the left) do not seem

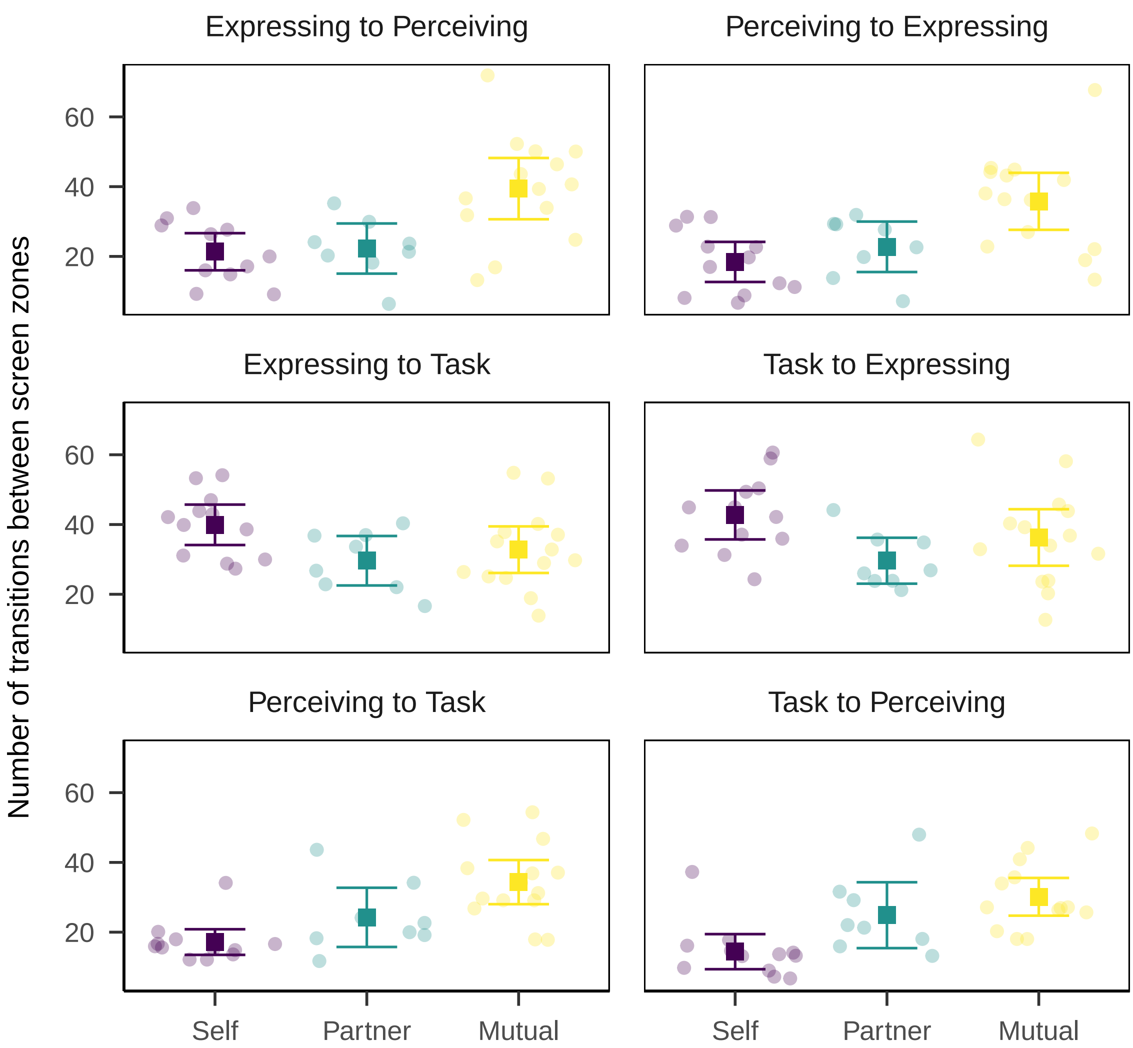


Figure 7.8: Number of transitions between Areas of Interest (AOI) on the interface. Transitions aggregated for $N = 34$ participants. Bars represent 95% confidence intervals.

to be more frequent compared to the *Partner-Oriented* condition. This may be relevant because it could rule out the possibility that a difference between the *Partner-Oriented* and *Mutual-Modeling* condition may be due simply to the fact that, in the *Mutual-Modeling* condition, participants only seek confirmation of what they have expressed, since this phenomenon is not available in the *Partner-Oriented* condition. It is therefore more likely that participants in the *Mutual-Modeling* condition seek the perceiving-monitoring zone for the emotional information about the partner.

Finally, the *Partner-Oriented* condition seems *stuck in the middle*. Results for this group are difficult to assess due to the small sample and the great inter-individual variance that is also present in the planned analysis. As a rule of thumb interpretation, the *Partner-Oriented* condition seems to go hand-in-hand with the *Self-Centered* condition in transitions between the *expressing-displaying* zone and the *perceiving-monitoring* zone (first row); and with the *Mutual-Modeling* condition in the other transitions (second and third rows).

In an attempt to figure out whether this kind of analysis may be used in a more structured manner, a multilevel linear model, also known as mixed linear model (D. Bates, Mächler, Bolker, & Walker, 2014; Kuznetsova, Brockhoff, & Christensen, 2017; West et al., 2015), was fitted to the data at hand using the mixed function of the Afex (Singmann et al., 2020; Singmann & Kellen, 2020) R package version 1.1.1. The model was fitted in the following terms: the number of transitions per participant for each possible path represented the outcome variable; the type of transition and the interface of the EAT (*i.e.* the experimental condition) were considered as fixed factors, with an interaction between the two; the participant was used as a random intercept to account for the non-independence of observations. A more complex model could have been more interesting, but hardly feasible due to the small number of participants (D. Bates, Kliegl, Vasishth, & Baayen, 2018).

A Type III Analysis of Variance of the multilevel linear model confirms effects of both individual factors and the interaction. Results are depicted in Table 7.6 using Kenward-Roger approximation for computing the p-value (Luke, 2017).

Table 7.6: Results of a Type III ANOVA on the fitted multilevel linear model

	num Df	den Df	F	Pr(>F)
group	2	31	5.5009	.009
transition	5	155	14.6194	<.001
group:transition	10	155	9.0331	<.001

Table 7.7 reports the pairwise comparisons between the three experimental conditions stratified by the bi-directional path of the transition. Differences in the pairwise comparisons can be detected between *Self-Centered* vs *Mutual-Modeling* and *Partner-Oriented* vs *Mutual-Modeling* in the transitions between expressing-displaying and perceiving-monitoring. In the four comparisons, the more socially-oriented interface obtained more transitions, in both directions, than the less socially-oriented one, corroborating the assumption that participants make use of the emotional information about the partner as a reference.

In the transitions between expressing-monitoring and the task, only a difference for the path going from the task to the expression-displaying zone was detected, with more transitions in the *Partner-Oriented* than in the *Self-Centered* interface. The effect is nevertheless not corroborated by any other comparison in the same transition path.

Finally, in the transitions between perceiving and the task, differences were detected between the *Self-Centered* and the *Mutual-Modeling* interfaces, with the *Mutual-Modeling* interface yielding more transitions in both directions. These results may support the role of emotional awareness as instrumental information directly related to the task at hand, but are not corroborated by a difference between the *Self-Centered* and the *Partner-Oriented* interfaces.

Table 7.7: Comparisons between the groups stratified by the path of the transitions. The Kenward-Roger approximation for the degrees of freedom is adopted and p-values are adjusted using the Tukey method for comparing a family of 3 estimates.

Comparison	Est.	SE	df	t.ratio	p.value
Expressing to Perceiving					
Self - Partner	-0.91667	4.8716	89.683	-0.18817	.981
Self - Mutual	-18.09524	4.1988	89.683	-4.30963	<.001
Partner - Mutual	-17.17857	4.7304	89.683	-3.63155	.001
Perceiving to Expressing					
Self - Partner	-4.33333	4.8716	89.683	-0.88951	.648
Self - Mutual	-17.36905	4.1988	89.683	-4.13668	<.001
Partner - Mutual	-13.03571	4.7304	89.683	-2.75575	.019
Expressing to Task					
Self - Partner	10.29167	4.8716	89.683	2.11258	.093
Self - Mutual	7.13095	4.1988	89.683	1.69833	.211
Partner - Mutual	-3.16071	4.7304	89.683	-0.66818	.783
Task to Expressing					
Self - Partner	13.12500	4.8716	89.683	2.69419	.023
Self - Mutual	6.46429	4.1988	89.683	1.53956	.277
Partner - Mutual	-6.66071	4.7304	89.683	-1.40808	.341
Perceiving to Task					
Self - Partner	-7.08333	4.8716	89.683	-1.45400	.318
Self - Mutual	-17.19048	4.1988	89.683	-4.09415	<.001
Partner - Mutual	-10.10714	4.7304	89.683	-2.13665	.088
Task to Perceiving					
Self - Partner	-10.45833	4.8716	89.683	-2.14680	.086
Self - Mutual	-15.72619	4.1988	89.683	-3.74541	<.001
Partner - Mutual	-5.26786	4.7304	89.683	-1.11363	.508

All things considering, transitions may represent a more interesting measure of the perceiving-monitoring function of emotional awareness compared to information seeking and information processing adopted in the planned analyses. Even considering the shortcomings (e.g. transitions interrupted by the use of the keyboard), transitions provide a more *dynamic* outlook on how emotional information is integrated into the task. The concept of transition may even be pushed further by measuring at which moment the transition has occurred, which would provide useful information on the interest – but also the potential distraction – of moment-to-moment awareness. For instance, it would be possible to assess whether learners look at the emotions expressed by the partner as soon as they appear on the interface, or if they wait idle period in the task. In this regard, the next analysis focus precisely on the moment-to-moment perspective, but with respect to emotion expression-displaying.

7.5.2 Emotions and Time: Evaluating the Purpose of Moment-to-Moment Awareness in Expressing Emotions

One of the main tenets of the present contribution is the advantage of moment-to-moment emotional awareness. Some exploratory analyses on the sample were performed in order to check to what extent the moment-to-moment feature has been exploited with respect to the expression of emotions.

Cognitive Evaluation Over Time

Congruently with appraisal theories of emotions – which state that it is the evaluation one does of the situation and not the situation *per se* that elicits the emotion – it is worth checking for the emergence of a pattern in the appraisals of *Valence* and *Control/Power* over time. Since all participants were exposed to the same stimuli (except participants in the *Self-Centered* condition, who did not see the emotions of the simulated partner), a clear pattern in the evaluation of the two criteria would not be congruent with appraisal theories. Figure 7.9 shows all the $N = 483$ emotions that have been expressed by all the participants over the 20 minutes of the task, organised by appraisal dimensions and stratified by experimental conditions. Each gray dot represents an emotion expressed by the participant, which are united by a gray line that forms an appraisal profile over time for each participant. A Locally Estimated Scatterplot Smoothing (LOESS) – that is, non-parametric curve that best fit the empirical data (Jacoby, 2000) – is superposed to the appraisal profiles for both appraisal dimensions, stratified by experimental condition.

The graphic indicates that there is indeed great variation in the appraisal profiles of each participant, with gray lines and dots spanning the whole range of the appraisal rating. When put together with the smoother, though, the patterns tend to overlap between appraisal dimensions, which corroborates the lack of orthogonality between Valence and Control/Power already observed with the Geneva Emotion Wheel and during the usability test of the Dynamic Emotion Wheel (see Sections 4.2.3 and 4.3.4).

With respect to the experimental conditions more specifically, the *less neutral* smoother corresponds to the Mutual-Modeling condition, for which both appraisal dimensions diminish over time. The data certainly does not allow any solid interpretation, but it is nevertheless worth noting that this kind of evolution could be investigated more thoroughly, for instance through the use fo spline regression (McElreath, 2020), which can account for the non-linearity of the process. In this way, it would be possible to test whether appraisal profiles tend to vary according to specific factors, for instance the emotions expressed by the partner.

Subjective Feelings Over Time

The same analysis can be conducted with respect to the expression of subjective feelings over time. Figure 7.10 plots the evolution of the expression of the 20 subjective feelings – which are part of the underlying EATMINT affective space used in the study – aggregated

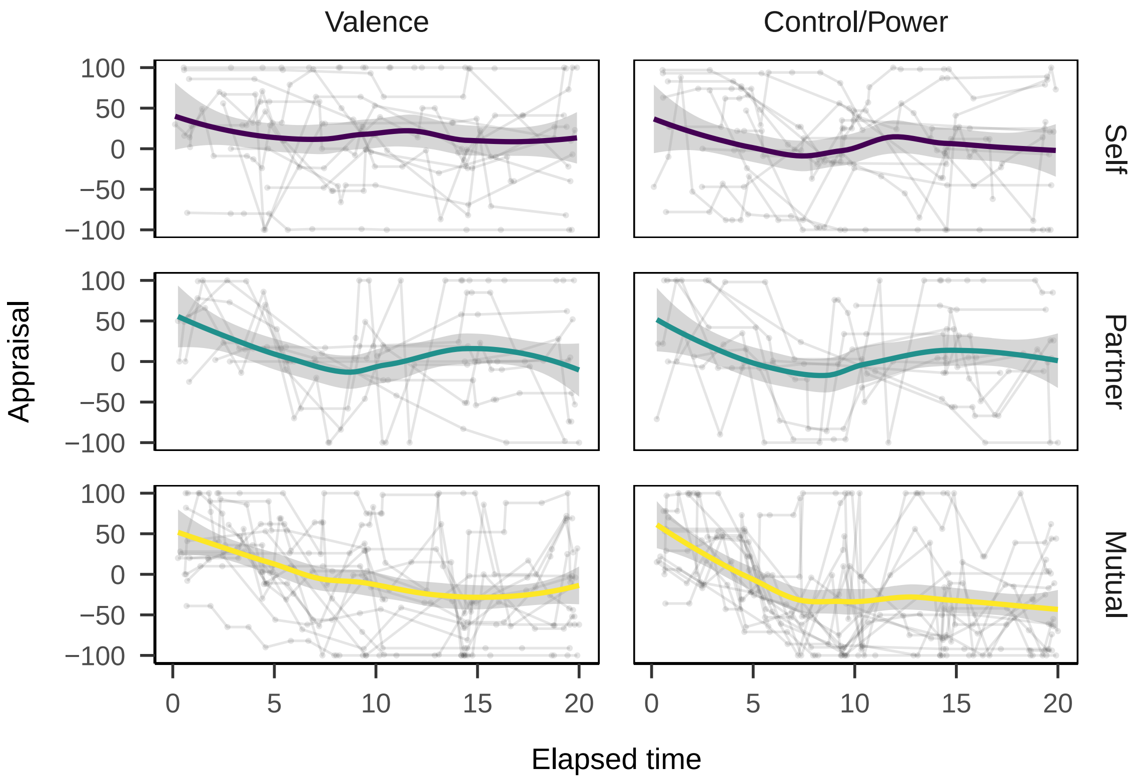


Figure 7.9: Evolution of the appraisal dimensions over time with a LOESS smoother ($N = 35$).

for all $N = 35$ participants. The observations are stratified per condition. In order to reduce the space of the graph, the legend for each condition has been omitted, but the colors are congruent with previous graphs.

Not considering the feelings that have been expressed only a few times (*e.g.*, *envious* or *disgusted*), most of the subjective feelings have been expressed rather uniformly over the 20 minutes of the task. Interesting exceptions are the feelings *bored* and *frustrated* that only starts around 5 minutes into the task – that is, around the end of the first enigma – which may be due to the repetitive nature of the task for boredom, and the increasing difficulty of the enigmas for frustration.

Finally, the overall small sample size combined with the unbalanced number of participants for experimental condition imply caution even on superficial interpretations about the effect of the interface. It is nevertheless worth noting how participants felt often *relieved* or *satisfied* in the *Mutual-Modeling* condition, but not in the *Self-Centered* or *Partner-Oriented* condition; or that the *empathetic* subjective feeling was expressed in the *Mutual-Modeling* and *Partner-Oriented* condition, but not in the *Self-Centered*. With a greater number of participants, it would be interesting to perform this kind of stratification in a more systematic way.

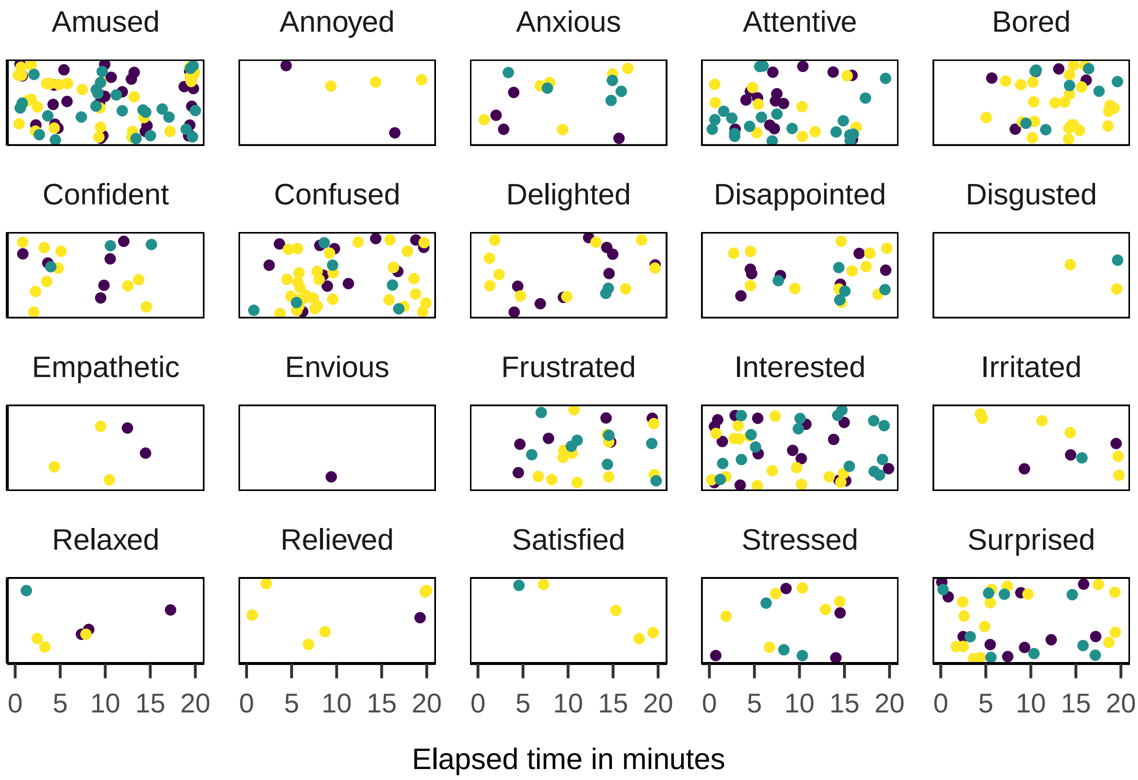


Figure 7.10: Expression of the subjective feelings. $N = 458$ emotions (out of 483) whose subjective feeling belongs to the underlying affective space used by the DEW, aggregated for the $N = 35$ participants. (Legend omitted for reducing space, see previous graphs.)

Expression of Emotions During the Different Parts of the Problem-Solving Task

Finally, a major conjecture that can be proposed with respect to the moment-to-moment feature of an EAT is that the emotional information is directly integrated within the task at hand. Since every enigma was divided in three sub-activities (reading, solving, and assessing the solution), it is worth investigating the number of emotions that have been expressed on each of the three type of activities. In fact, it may be assumed that the solving part of the task is the one that requires more interaction with the partner, since participants were supposed to base their solution on each others reasoning. In this sense, the number of emotions shared during the solving part of the enigma would suggest that the emotional information is considered instrumental to the task at hand.

To assess this claim, the number of emotions expressed by each participant in each of the three parts of the task was averaged across the four enigmas. Since the three sub-activities did not last the same amount of time (40 seconds for reading, 3:20 seconds for solving, and 60 seconds for assessing the solution), each average was also projected as if the whole task of 20 minutes was potentially constituted of only the interested part. Figure **??** depicts the observed as well as the projected number of expressed emotions stratified by task sub-parts and experimental conditions. The horizontal gray line corresponds to the observed mean of expressed emotions, that is $M = 13.8$.

Figure 7.11: Observed and projected number of expressed emotions depending on the type of activity during the task (reading, solving or assessing the solution). The gray line corresponds to the observed overall mean of $M = 13.8$.

The data suggest that, proportionally, participants expressed more emotions while reading the enigma and, in particular, assessing the solution, whereas fewer emotion were expressed during the problem-solving task. This is most probably due to the fact that solving the problem needs more cognitive effort and also direct manipulation on the keyboard to write the reasoning compared to the passive situation of reading the enigma and finding out the solution. At the same time, it is interesting to note that the number of emotion expressed is more or less consistent across experimental conditions, which corroborates the interest for an intra-individual perspective already mentioned above, in particular when facing the problem and discovering the solution – that is, two situations which are novel and relevant for the person (Scherer, 2013b).

7.5.3 Internal Meta-Analysis on Task Indicators

Taking advantage of the fact that the same task was adopted, under similar conditions, in the usability test of the DEW (Fritz, 2015), small internal meta-analyses (Goh et al., 2016) were performed on the point-estimate means for the three dependent variables adopted in the current contribution. The interest of the meta-analyses is to provide a better assessment of the performance-based indicators of the use of the EAT, for instance as reference for future studies[1].

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

Each meta-analysis has been conducted using the R meta package version 5.2.0, adopting the inverse of the variance weighting mechanism to account for differences in the sample size of the two internal studies (Borenstein, 2009). Results both for the fixed and the random models (using the DerSimonian-Laird estimator for tau^2) are provided.

Expressing Emotions Internal Meta-Analysis

The meta-analysis on the expression of emotions has been conducted on all the retained participants for both studies, since, even for participants in the *Self-Centered* condition, the overall situation in which participants have expressed their emotions are sufficiently close for an internal meta-analytic purpose. Consequently, the sample size are of $N = 14$ in Fritz (2015) and of $N = 35$ in this contribution. Results, depicted in Figure 7.12, assess an estimated mean of 14.92 [13.39; 16.46] expressed emotions for the fixed effect model, and of 15.34 [12.07; 18.60] for the random effect model. The meta-analysis highlights the presence of considerable heterogeneity in the expression of emotions ($\tau^2 = 10.29$; $\tau = 3.21$; I^2 = 82.0% [23.9%; 95.7%]; H = 2.36 [1.15; 4.84]; $\chi^2 = 5.55$ $p = .018$). This may suggest that the different conditions of the two studies may have played a role in inflating the number of emotions expressed in Fritz (2015), where participants were explicitly asked, if possible, to express at least one emotion in each phase of the 4 enigmas (*i.e.* which would amount to 12). On the contrary, in this contribution they did not receive any guidance as of the number of emotion to express. This may be interpreted as a warning about the importance of being careful in framing how the expression of emotional information is prompted, even if the inflation of the number of emotions may not be necessarily accounted by *forced* emotions, that is, emotional episodes that are not *really* felt, but nevertheless reported. It may also be the case that prompting for emotional expression may ease participants into expressing their emotions, something they could be less prone to do otherwise.

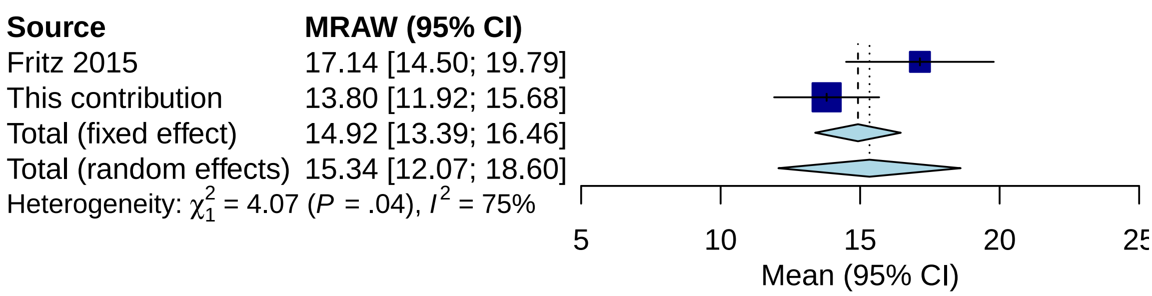


Figure 7.12: Internal meta-analysis of the number of emotions expressed in the experimental task.

Information Processing Internal Meta-Analysis

The internal meta-analysis on information processing has been conducted using only the participants retained for the eye-tracking analysis ($N = 12$) in Fritz (2015), and only participants in the *Partner-Oriented* and *Mutual-Modeling* conditions ($N = 23$) in this contribution, for the interface in these situations is identical or at least very similar with respect to the *social* information shared. Results, depicted in Figure 7.13, assess an estimated mean of 60.14 [51.15; 69.13] total visit duration, in seconds, for the fixed effect model, and of 60.14 [51.15; 69.13] for the random effect model. The meta-analysis does not detect heterogeneity between studies (τ^2 = 40.81; τ = 6.39; I^2 = 45.2%; H = 1.35; χ^2 = 1.82 p = .177), which may indicate that the time spent at looking at emotional information could be determined by a balance between the primary problem-solving activity and the sustaining emotional awareness. The point estimate of total visit duration being around 1 minute over the 20 minutes of the task, it corresponds to a proportion of 5% of the total time.

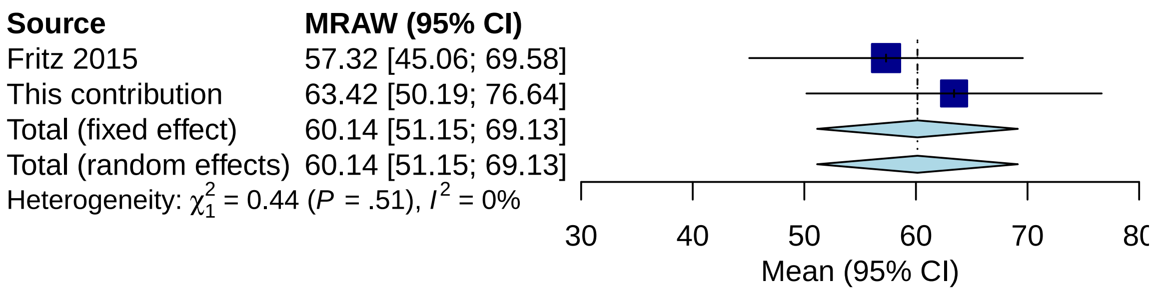


Figure 7.13: Internal meta-analysis of the time spent at processing emotional information available on screen.

Information Seeking Internal Meta-Analysis

With respect to emotional information seeking, the internal meta-analyses comprise the same samples as for information processing, that is $N = 12$ in Fritz (2015) and $N = 23$ in this contribution. Results, depicted in Figure 7.14, assess an estimated mean of 81.63 [71.59; 91.67] number of visits for the fixed effect model, and of 81.63 [71.59; 91.67] for the random effect model. The meta-analysis does not detect heterogeneity between studies (τ^2 = 0; τ = 0; I^2 = 0%; H = 1; χ^2 = .56 p = .453), which is nevertheless rather due to the huge variability within studies rather than homogeneity between studies. In future studies, the number of transitions between AOI could represent a more informative measure for emotional information seeking.

7.6 Discussion

The planned analyses in this experiment aimed at investigating whether a different use of, and access to emotional information determine differences in the concrete use of an EAT during a computer-mediated collaborative task, which has been simulated in order to control from other sources of emotional information outside the manipulated interface of the EAT. The technical issues, due to articulated settings between the material and the simu-

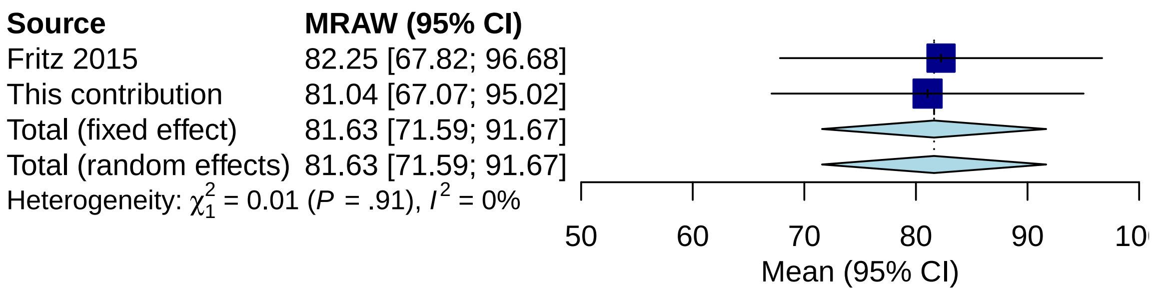


Figure 7.14: Internal meta-analysis of the number of times emotional information has been visited on screen.

lation implicated in the experiment, have nevertheless reduced an already limited sample size. Furthermore, the inter-individual differences in all dependent variables entail wide confidence intervals, whose source can be traced back to the many processes implicated in the task. Some of them may not be directly inherent to a genuine interest (or disinterest) in emotional awareness, and may have potentially influenced participants' capacity in conveying and taking into account emotional awareness beyond their intentions. For instance, participants had to coordinate multiple functions, both cognitively and practically (e.g. writing on a keyboard, manipulating the EAT, etc.), under specific time constraints. Participants with less dexterity in writing at the keyboard or manipulating the interface may have found less time to dedicate to the EAT even if they were willing to. The small sample size cannot guarantee that these individual differences are sufficiently balanced by the randomized trial. Even in the presence of detectable effects, thus, the assessment of their relevance in terms of *practical* consequences is limited: their size is inherently high due to the small sample and they should not even be taken as reliable benchmarks for future studies (Albers & Lakens, 2018).

On the other hand, the controlled environment in which the *performance-based* measures have been obtained make them worth of interest in assessing to what extent the presence of an EAT serves as an affordance in conveying and taking notice of emotional awareness during a computer-mediated collaborative task. Eye-tracking measures, in particular, may be considered spontaneous reactions occurring to some extent even beyond participants' top-down control (Jacob & Karn, 2003; Poole & Ball, 2005). The discussion of the obtained results may thus contribute to sketch a more defined outlook of the use of an EAT and provide cues for further hypotheses worth investigating or shortcomings to be taken into account in future studies. At the same time, corollary analyses can also be implemented in a more general assessment of the EAT with respect to its fundamental features: self-report, emotion structure injected into the tool, and moment-to-moment availability of emotional awareness.

7.6.1 Emotion Expression Seems Viable Also From a Self-Centered Perspective

In the first hypothesis, it has been posited that learners expression of emotions through the EAT varies depending on what use would be made of them, and what emotional information they have access to through the interface. More precisely, it has been stated that participants in the *Self-Centered* condition would express fewer emotions, compared to the other, more socially-oriented, conditions. It has also been posited that participants in the *Mutual-Modeling* condition would express more emotions than in the *Partner-Oriented* condition by virtue of an additional prompt in social sharing due to the direct and persistent comparison between one's own emotions and that of the partner.

This hypothesis could not be supported, with results that also revealed a violation of assumption about normality of residuals in the ANOVA model (Hoekstra et al., 2012). The number of emotion expressed may in fact be considered a tricky measure, since, depending on the specific number of emotion expected, it may be considered a countable measure, thus closer to a Poisson sampling distribution rather than a Gaussian one. Notwithstanding, the usability test conducted in similar conditions yielded an average of emotions around 17, and this experiment yielded an average close to 14, with neither of the measures that was zero-inflated as it is often the case with *pure* countable measures. As a consequence, the use of ANOVA could be warranted. The non-normality of residuals in this specific case may rather be due, thus, to an intrinsic inter-individual difference towards expressing emotions, which can result from a mixture of dispositional stances (Scherer, 2010a, 2021a) and technical skills in manipulating the interface. If this is the case, then, small effects accountable to the interface should be expected. As a consequence, researchers planning for effects of contextual/external factors on the expression of emotion shall probably aim at a considerable sample size to discern the signal from the inter-individual noise.

It may thus be interesting to reverse the perspective and, rather than seeking for differences, highlight how participants in all conditions expressed on average $M = 13.80$ ($SD = 5.68$) emotions, that is more than 1 emotion every 2 minutes. In particular, participants in the *Self-Centered* condition expressed on average $M = 12.50$ ($SD = 5.30$) emotions despite knowing they were the only recipient of the information. This result could be considered, in principle at least, as supportive of the intra-personal function of emotional awareness. The presence of an EAT could have increased emotion alertness as a first step towards emotional expression, which may then be used as a form of (implicit) emotion regulation (Lieberman et al., 2007; Torre & Lieberman, 2018). On the other hand, though, this result may also be explained by side effects of the experimental task. For instance, this number may be inflated by task compliance, since the overall experimental setting was overtly aimed at expressing emotions. Furthermore, the characteristics of the experimental task, whose timing is fixed and not determined by the participants' actions, may also have pushed participants to express emotions to fill idle time between enigmas or part of the task within each enigma, rather than for an urge to express and regulate their emotions. This seems corroborated by the corollary analysis comparing the number of emotions expressed during the three different parts of each enigma: reading, solving, and assessing the solution. Participants expressed proportionally more emotions while reading the enigma and assessing the solution compared to the problem-solving task, in which emotional awareness could be directly integrated into collaborative processes such as transactivity (Gaëlle Molinari, Chanel, et al., 2013; GaÃÆ'NANAÃNAÂlle Molinari, Sangin, Dillenbourg, & NÃÂss, 2009; Mirweis Sangin et al., 2011) or the resolution of socio-cognitive tensions (Andriessen et al., 2011). In other words, it would seem that participants felt the need to express how they felt from an holistic perspective. Even in the conditions with emotional information that could potentially be used for partner- and mutual-modeling, participants did not seem to feel the urge to convey more emotions to the partner. On the other hand, when participants were more free to assess the situation, they indeed expressed how they felt. The time constraint imposed and the simulated nature of the task require nevertheless caution on this kind of inference. Especially the time constraint introduces a clear bias in the rhythm of tension and relaxation that could be beneficial to producing and consuming emotional awareness even while performing an active task.

The moment-to-moment possibility of expressing emotion needs, thus, a more ecological situation for a thorough assessment, compared to the controlled environment adopted in this experiment. On the other hand, the strict environment allowed to expose all participants to exactly the same condition, except for the manipulated interface. In this regard, the assessment of the appraisal dimensions have shown that, congruently with appraisal tenets, participants evaluated the very same situation in different ways, especially through the use of the sliders. This corroborates the interest for implementing a dimensional approach into the tool, especially from an appraisal-driven perspective. On the other hand, the evaluation on the two sliders seem to overlap, a general problem that will be assessed more thoroughly in Chapter 9.

All things considering, thus, the present contribution conveys limited and mixed evidence with respect to the interest for expressing emotions during a computer-mediated collaborative task as a function of the use of and access to emotional information. The expressing-displaying function is also the one more influenced by the task at hand. Further investigation on the subject should thus not only increase the sample size, but also implement a real collaboration with more relaxed time constraints. The dynamics within each group or dyad of participants could be accounted for by multi-level models, and therefore a strict controlled environment can be deferred in favor of a more ecological setting. Further investigation may also plan for a sample size determined on an equivalence test, and perform both tests (Fidler, Thorn, Barnett, Kambouris, & Kruger, 2018; Lakens, 2017): one in search of a difference, and one that attempts ar ruling out an effect of the use of, and access to emotional information. With respect to the latter, potential limits for the same kind of task – but with real collaboration – could be set in raw units at +/- 3 emotions, which roughly corresponds to half a standard deviation from the internal meta-analysis.

7.6.2 Emotional Information Seeking and Processing Seem Corroborated by a Socially Oriented Perspective

In the second hypothesis, it has been posited that learners would seek and process the emotional information available on screen depending on whose emotions were available on screen and whether a direct and persistent comparison was enhanced or not. More precisely, it has been stated that participants in the *Self-Centered* condition would seek less often and process for shorter time the information available through the perceiving-monitoring part of the interface compared to the *Partner-Oriented* and *Mutual-Modeling* conditions. It has also been posited that participants in the *Mutual-Modeling* condition would seek information more often, and process it longer compared to the *Partner-Oriented* condition due to the increased interest enhanced by direct and persistent comparison of emotional information.

Results corroborate the presence of an overall effect of the interface both on emotional information seeking and processing. For information seeking, the experimental condition yielded a generalized effect size (Olejnik & Algina, 2003) of $\hat{\eta}^2_G = .317$, 90% CI $[.092, .490]$, whereas for information processing the effect consisted in $\hat{\eta}^2_G = .281$, 90% CI $[.064, .457]$. In both cases, thus, the experimental condition seems to account for a considerable amount of variation in the perceiving-monitoring function of emotional awareness, even though the confidence intervals are once again wide enough to include a small overall effect.

On a more fine-grained level, the differences between conditions only partially corroborated the directional hypothesis. Differences were detected, both for information seeking and processing, only in pairwise comparisons between *Self-Centered* vs. *Partner-Oriented* and *Self-Centered* vs. *Mutual-Modeling*, but not between *Partner-Oriented* vs. *Mutual-Modeling*. The difference for seeking emotional information are wider between the *Self-Centered* and the *Mutual-Model*, with en estimation of around 44 more visits in total, and confidence intervals spanning between 15 and 72 total visits. The difference between the *Self-Centered* and the *Partner-Oriented* conditions is lower and more uncertain, with an estimation of around 35 more visits in total, but confidence intervals between around 2 to 67 visits. On the contrary, for emotional information processing, the difference is wider between the *Self-Centered* and the *Partner-Oriented* conditions, with an estimation of around 40 additional seconds spent processing the graphical representation of emotions, and confidence intervals between around 10 and 71 seconds. The difference between the *Self-Centered* and *Mutual-Modeling* conditions was instead of around 32 additional seconds, with confidence interval between around 4 to 59 seconds. The effect seems therefore consistent for both seeking and processing emotional information, but the estimation too uncertain to determine whether the difference could be of any practical value.

Furthermore, the lack of a detectable difference between the *Self-Centered* and the more socially-oriented interfaces would have undermined the usefulness of an EAT, whereas its presence may be explained as a *simple* novelty effect. The *Partner-Oriented* and *Mutual-Modeling* condition convey information that the learner does not know, whereas in the *Self-Centered* condition the emotions are just a reminder of what the learner should already know. In this regard, compared to emotion expression for which the estimated mean of 12.5 emotions was considerable, the estimation for seeking and processing emotion information leave more doubts about potential intra-personal meaning-making extrapolated from disposing of one's own emotions on screen, even if the representation facilitates the perception of the emotional experience evolution. The more socially-oriented interfaces also presented a word cloud, which may have been perceived as a more interesting representation compared to an additional graph (Pinker, 1990).

The comparison between the *Partner-Oriented* and *Mutual-Modeling* interfaces has, on the other hand, deeper implications with respect to the *raison d'être* of an EAT, and it also represents a more severe test given that the two conditions are more similar (Mayo, 2018). The lack of a discernible effect between the two social-oriented conditions, if confirmed with an equivalence test with increased power, may suggest that there is no additional value conveyed by direct and persistent comparison available on screen. But the phenomenon could also be explained by the fact that the additional value is gained without the need for further information seeking and processing. That is, participants in the *Mutual-Modeling* condition were able to compare their own and the partner emotion thorough the interface without having to look at the perceiving-monitoring zone of the EAT more often or longer, because they could get more information with the same effort. The eye-tracking measures alone cannot unravel whether the lack of a discernible effect is a positive or negative outcome with respect to the social-oriented hypothesis. In future studies, the hypothesis should first be corroborated by an equivalence test (Lakens, Scheel, & Isager, 2018), and then also be assessed with the aid of self-reported measures about the perceived usefulness of direct and persistent comparison of learners' emotional experience.

On the other hand, the use of transitions as a more dynamic indicator to measure the perusing of emotional information corroborates an overall increased activity within the *Mutual-Modeling* condition where the perceiving-monitoring zone of the tool is implicated. Even though the difference in the exploratory model are not always statistically significant, it is possible to outline some patterns. For instance, the gaze-path between the expressing-displaying and the perceiving-monitoring part of the interface counted around 17 transitions forth, and 13 more transitions back in the *Mutual-Modeling* compared to the *Partner-Oriented* condition. This may suggest that the direct and persistent comparison was taken into account before and/or after expressing an emotion, which can be interpreted as a mutual-modeling activity in itself.

The more interesting path between the perceiving-monitoring part and that of the task, on the other hand, indicates more than 15 transitions (back and forth) in the *Mutual-Modeling* condition compared to the *Self-Centered*. This may also be considered an indicator of some mutual-modeling activity, even though this phenomenon is not corroborated by differences implicating the *Partner-Oriented* condition, which do not reach the significance threshold.

All things considering, thus, the experiment, even without stating any stable estimation on the effects, corroborates a favorable outlook towards the two more socially-oriented interfaces with respect to emotional information seeking and processing. Having the emotion of someone else on the interface produce at least the curiosity to seek that information and also process it, whether valuable meaning-making is extrapolated or not. Emotions of the partner also increase the number of transitions in the *Mutual-Modeling* condition where the perceiving-monitoring zone is implicated. This is consistent with the idea that emotionally-charged stimuli have a privileged access to attention (Brosch et al., 2010; Pool et al., 2015). It also suggests that despite the limitation of the graphical representation of emotion available through the toolbox, those are considered interesting enough to be sought and processed. The development of better visualizations of the emotional information should therefore considered more confidently (Berset, 2018; Derick et al., 2017; Fritz, 2016b; Leony et al., 2013). At the same time, these measures can be assessed as a more hectic attitude towards the overall interface, with participants switching their gaze more erratically back and forth between zones. Under this assumption, the EAT would represent rather a distraction, considering once again the salience of emotional stimuli in calling for attention (Brosch et al., 2010; Brosch et al., 2011; Pool et al., 2015). In this regard, it may be useful to perform a more detailed analysis on when the participants' gaze has transitioned into the perceiving-monitoring zone, for instance by using the expression of emotions from the partner as cues after which transitions should be looked for in a short period of time. This analysis has not been performed since the synchronization between the timing of the eye-tracker software and the toolbox was difficult to achieve. It will nevertheless be possible to synchronize the two timer in future studies.

7.7 Conclusion

This chapter presented a detailed illustration of an empirical study investigating whether a different use of, and access to emotional information expressed and available through an EAT had an effect on the actual use of the EAT itself. $N = 48$ participants, then reduced to $N = 35$ mostly for technical reasons, were randomly assigned to three different interfaces of the EAT – namely a *Self-Centered*, a *Partner-Oriented*, and a *Mutual-Modeling* interface – which varied on how socially-oriented each interface was. The main assumption underlying the empirical investigation stated that the more socially oriented interface would yield a greater use of the EAT in terms of emotion expressed and emotional information seeking and processing. This assumption was not corroborated for the number of emotion expressed, and was only partially corroborated for emotional information seeking and processing, for which participants in the *Partner-Oriented* and *Mutual-Modeling* conditions sought and processed emotional information to a gretaer extent compare to the *Self-Centered* condition, but without a detectable difference between the two more socially-oriented conditions.

Despite the hypotheses of the study being rejected or only partially corroborated, most of the performance-based indicators about the use of the EAT were congruent with the main assumption. In most measures, the *Mutual-Modeling* interface yielded the greater use of the EAT, followed by the *Partner-Oriented*. Congruently with a inter-personal perspective on emotional expression (Parkinson, 1996; Parkinson & Manstead, 2015; Rimé, 2009; Van Kleef, 2009, 2018; Van Kleef et al., 2011), these results seem to corroborate the usefulness of an EAT as an affordance to share emotion with a partner and take the partner's emotions into account during a computer-mediated collaborative task. Participants seemed to show a genuine interest in seeking and processing emotional information available about the partner, and knowing that their emotions would be conveyed to the partner did not stop participants to express them. The presence of the partner's emotions on the EAT interface also resulted in a more *dynamic* gaze-path, with more transitions between the part of the interface dedicated to the task, and that dedicated to the monitoring-perceiving function of awareness. This fact seems to corroborate the usefulness of providing moment-to-moment emotional awareness, since the emotional information may be truly integrated as instrumental information to the task at hand (Buder, 2011; Dourish & Bellotti, 1992). A word of caution is nevertheless in order, since the method by which transitions have been computed is not externally validated yet. Whether and when it will, transitions could represent a more adequate measure of integrated and dynamic information seeking and processing compared to the *static* number of visits and seconds spent inside an Area of Interest used in the directional hypotheses of the study. All things considering, thus, a moderate optimism is warranted about the usefulness of an EAT during a computer-mediated collaborative task. Despite the fact that it does not directly provide information about the content space of the task (Janssen & Bodemer, 2013), sharing emotions could nevertheless sustain the mutual-modeling activity by which learners build and update a holistic representation of their partner in the collaboration (Dillenbourg et al., 2016; Gaëlle Molinari et al., 2009; Mirweis Sangin et al., 2011).

At the same time, participants in the *Self-Centered* condition also seemed to harness the presence of the EAT, which is congruent with an intra-personal usefulness in expressing emotions (Lieberman et al., 2007; Lieberman, 2019a; Torre & Lieberman, 2018). Even though participants in this condition knew beforehand that they were the exclusive sender and receiver of the emotional information, they still expressed emotions as well as sought and processed their own emotional information available on the EAT. This fact suggests that the presence of an EAT may prompt learners to inquiry about their own emotional state, appraising the situation and seeking for the congruent subjective feeling elicited by the circumstances (Boehner et al., 2007; Grandjean et al., 2008; Scherer, 2009b).

7.7.1 Limitations and Future Development

The present contribution adopted a controlled environment in order to expose every participant to the same stimuli, except for the randomly assigned interface. The use of a simulated partner limited the inter-personal communication that would be normally available in a real collaborative setting. For the purposes of the study, a distinction between cooperation (roughly, doing the same thing with limited interdependence) and collaboration (integrating efforts into a common outcome) was not of primary concern. Nevertheless, this certainly represents a limitation to the generalization of the obtained results to a more articulated communication flow, which could overlap with the emotional information expressed through the EAT. For instance, learners could have used the text area dedicated to their reasoning to inject circumstantial information such as *I don't understand* or *I don't agree with you*, which would have conveyed social, cognitive and emotional information (Derks et al., 2008). A focal question avoided by the present experimental setting is therefore whether participants would still have used the EAT the way they did, were they allowed to convey emotional information through the content space of the joint problem-solving activity. The task at hand, though, can be easily extended to a real collaboration between two participants, and therefore it would be possible to investigate the matter in a way that can be directly related to the results of this contribution. Rather than for its result, this experiment may be more relevant for its design: testing different versions of an EAT, where differences have also theoretical or pedagogical implications. The same design can be progressively enriched, first by the same task, but with real collaboration, and then by introducing tasks that are closer to authentic Computer-Supported Collaborative Learning activities.

With respect to an hypothetical next step consisting in a direct replication with real collaboration, the comparison with present results would have to take the difference in time and setting into account. In this regard, a possible solution is to retain a ratio between the duration of the joint-problem solving task and the performance-based measures of the use of the EAT. For example, participants in a dyad that solved the four enigmas in 12 minutes and expressed 18 emotions for one participant and 24 for the second would have a ratio of 1.5 and 2 respectively. Taking into account the temporal dimension would also add another element of interest: investigate whether the ration holds constant across different durations of the joint problem-solving task, or it yields a moderation effect.

Furthermore, the focal question of whether learners would use the EAT having a more thorough channel of communication could also be assessed by allowing learners to display or not the EAT on their screen. This choice may also be instrumental in investigating one of the main assumptions of the thesis, namely the interest for moment-to-moment emotional awareness: whether and when participants would decide to display the EAT may convey pivotal evidence about the usefulness of moment-to-moment emotional awareness compared, for instance, to a *scripting* strategy in which partners share their emotions at specific intervals outside the task.

7.7.2 Experiment's Overall Contribution

To sum up, in spite of the technical difficulties and the limited external validity, the experiment provided contributions that may be retained for future studies. First, breaking down the abstract model of emotional awareness can be a useful guideline to set up experiments. Second, the use of transitions between Areas of Interest can be maintained as a sensible, reliable and valid measure for a genuine interest in perusing emotional information through the EAT. Third, the specific measures provided by the toolbox can be analysed in different ways, empowering the investigation of emotional awareness.

7.7.3 Acknowledgments

The experimental phase of the project has been carried out by Stéphanie Perrier as part of her Master thesis (Perrier, 2017) that Mireille Bétrancourt and I co-directed.

Chapter 8

Emotional Awareness in Blended Learning: an Exploratory Analysis of the Use and Perception of an Emotion Awareness Tool Over Time

This chapter describes the second empirical contribution of the thesis, which investigates the longitudinal use of an EAT, in asynchronous and non-collaborative learning settings, made by two classes of a blended Master in education technology. As for the previous chapter, the presentation of the study follows the traditional structure of empirical contributions (Sollaci & Pereira, 2004) and therefore the outline of the chapter is not presented. The chapter starts directly by introducing the rationale of the study.

8.1 Study Rationale

Even though there is nowadays a consistent agreement in considering emotions as bounded episodes unfolding over relatively short time (Pekrun & Linnenbrink-Garcia, 2014b; Sander, 2013; Scherer, 2019b), the effects of emotional experiences also have consequences on the long term (Brackett, 2019; Pekrun et al., 2018; Rimé, 2005; Scherer, 2021a). As stated by appraisal theories of emotions, for instance, emotions are elicited by the evaluation of events considered of major concern for the organism, and that evaluation can over time present regular patterns, for which a person may be more prone to evaluate events according to stable perspectives (Scherer, 2021a). Emotions play also a prominent role in the formation of meaningful and stable relationships with others (Barsade & Knight, 2015; Rimé, 2005, 2009; E. R. Smith, Seger, & Mackie, 2007; Van Kleef & Fischer, 2015). As already mentioned in Chapter 1, there is also a growing consensus in learning and educational sciences to take a more holistic and integrated approach to students' emotions at various levels, from the everyday experience to the creation of curriculum that foster socio-affective competences

(Brackett, 2019; Henritius et al., 2019; Pekrun & Linnenbrink-Garcia, 2014a; Pekrun et al., 2018).

In computer-mediated learning environments, especially applied to asynchronous and remote settings, the affective experience play a prominent role in determining whether students manage to *keep on track.* For instance, students may still manage to develop a sense of belonging to a group in spite of physical distance, but they may also risk to feel isolated and progressively loose momentum and motivation to follow through Henritius et al. (2019). In this regard, there is a growing consensus in considering that learners should be given the opportunity to self-reflect on their emotional experience (Lavoué et al., 2020; Lavoué et al., 2017; Gaëlle Molinari et al., 2016; Ruiz et al., 2016), and at the same time project themselves socially and affectively into the computer-mediated learning environment as a means to enhance the *social presence* (Lowenthal & Snelson, 2017 for an overview, see also Section 1.3.3).

As a consequence, scholars have started to upscale the investigation of emotion in learning, by tracking the emotional experience of students over longer period of times (Lavoué et al., 2020; Lavoué et al., 2017; Gaëlle Molinari et al., 2016; Ruiz et al., 2016). For instance, Ruiz et al. (2016) introduced the Twelve Academic Emotion Model (TEAM) mentioned in Section 1.1.2 during two semesters in on-site courses in Computer Science classes, during which teachers asked participants to rate their emotional experience at specific moments through a computer-mediated interface. The overt aim of the practice was to stimulate self-reflection on emotional experience, which was supported by the possibility to monitor the emotional information through a series of graphical representations. The authors report mixed results about how the usefulness of this process was rated by participants, even if, all things considering, they depict a favorable outlook overall. Lavoué et al. (2020) followed 11 students at the University level for several weeks in the middle of the second semester. They provided students with a grid through which students could record their emotions from a list of 9 discrete emotions. Retrospective interviews were also conducted, which were focused on an appraisal-driven perspective. Students were asked about the causes and evaluation of the events they have been recording through the grid. Finally, as already mentioned in the related works of Section 2.2, Gaëlle Molinari et al. (2016) used the the EMORE-L self-report tool to track the emotional experience of students over 15 days, but exclusively from an intra-personal perspective. In this regard, participants expressed the wish to dispose also of information about the emotional information of others.

In the same vein, it would be worth investigating whether the EAT implemented in the thesis may serve as a viable instrument for students to self-reflect on their emotional experience over a longer period of time, and also assess whether the tool allows them to project themselves socially and affectively in a computer-mediated learning environment in an asynchronous and non-collaborative situation. Such a setting is particularly relevant, since it does not structurally demand social interaction between learners (Henritius et al., 2019; Jézégou, 2010; Kirschner et al., 2015; Kreijns et al., 2003). As a consequence, socio-affective interaction is not stimulated the instructional design *script.* The introduction of an EAT may therefore serve to accomplish this function.

In this regard, the present study adopts a longitudinal plan (Fitzmaurice et al., 2011) in which the use of the EAT is implemented in an ecological context of distance learning. The Master of Science in Learning and Teaching Technologies (MALTT) at Geneva University provides a blended learning program since more than 20 years. The planning divides each semester in three periods, in which a week of on-site classes is followed by 4-5 weeks of remote learning. The use of the EAT will be implemented in one of the courses of the Master, named *Sciences et Technologies de l'Information et de la Communication I* (STIC I), which covers introductory web programming and computational thinking (Fritz & Schneider, 2019). Students will use the EAT to express their emotions at any moment while they are working for the STIC I course – outside the on-site sessions – for which they will be implicated in individual assignements. The collection of data is therefore comparable to the Experience Sample Method [ESM; Csikszentmihalyi & Larson (2014)] or the Ecological Moment Assessment [EMA; Shiffman et al. (2008)] already used in distance learning situations (*e.g.*, Gaëlle Molinari et al. (2016), but see also Scollon et al. (2003) for a critical assessment of the method). There are nevertheless two important distinctions. First, contrary to some implementation of ESM or EMA, in which there is an external prompt that reminds participants of the recording activity, the use of the EAT is left to the spontaneous initiative of students, who can decide whether and when to use it based on the eliciting events during remote learning (Wheeler & Reis, 1991). In this way, the use of the tool remains closer to the purpose of an awareness tool. Second, the information will not only be collected, but also available both to the learner herself and to the other students of the class. In other words, all the passages of the abstract model of emotional awareness depicted in Section 2.5 are implicated in the process.

8.2 Research Questions

The overall settings in which the study takes place can therefore contribute to assess whether and, possibly, why the EAT is adopted in the first place, as well as the evolution of its use and perception of its usefulness over time. Given that the EAT has never been adopted before in longitudinal settings, though, the overall character of the study is overtly exploratory, without any hypothesis stated beforehand (Haeffel, 2022; Scheel et al., 2020b). The research questions addressed in this study can be grouped in five main topics, detailed hereafter.

8.2.1 Use of the EAT Over Time

A first topic of interest concerns, bluntly, whether the EAT is used in the first place. One thing is to dispose of a tool for a limited period of time, as in the experiment of the previous chapter, where novelty and curiosity can play a prominent role. Another is to sustain the effort of producing and perusing emotional information over a longer period of time, during which resources are dedicated to the coordination of different tasks (reading learning material, design a project, debug code, etc.). The subject is particularly relevant from an awareness tool perspective, where learner are bestowed with the responsibility of adopting the tool if and when they see fit, rather than being guided by the instructional design (Miller & Hadwin, 2015).

The interest of a longitudinal plan also relates to the possibility that the use changes over time. Novelty and curiosity can induce the initial adoption of the tool, but, at length, the EAT may progressively be dismissed. On the contrary, it may also be envisaged that an initial reluctance can be then steered in a more convinced use after acknowledging its benefits. Question 1 (Q1) can therefore be stated in the following terms: is the EAT adopted in the first place, and does its use evolve over time?

8.2.2 Perception of the Usefulness of the EAT Over Time

As highlighted in the previous experiment, which was mainly focused on the concreted use of the tool (i.e. the left-hand side of the abstract model), the meaning-making function of emotional awareness cannot be inferred exclusively from performance-based indicators. The assessment of the right-hand side of the abstract model can benefit from an assessment based on the perception learners have of the EAT's usefulness. As it is the case for the use, also the perception can vary over time, and the two may not be necessarily linked. One can stop using the EAT even though the perception of its usefulness may remain high, for instance because the learner has too many things to do and therefore needs to sacrifice a *side activity*. On the contrary, one can continue to use the tool for normative (task compliance) or social pressure (the colleagues use it), but without perceiving any real value.

On this topic, the contribution introduces a tentative scale, named Emotion Awareness Usefulness. The scale attempts to provide cues that accounts for the many facets of emotional awareness illustrated by the abstract model, including the right-hand side more tightly linked to meaning-making, which is more difficult to assess with pure performance-based indicators. The scale consists in seven dimensions – frequency, affordance, social presence, self-understanding, understanding others, self-other comparison, and self-regulation – derived from the literature and/or by the functions of emotional awareness. The scale is presented and discussed in the Material part of the Methods section below. The use of a validated usability scale, the System Usability Scale (SUS) by Brooke (1996), will also be used in this perspective. Question 2 (Q2) is stated as follows: what is the perception of usefulness of an EAT in asynchronous and non-collaborative settings, and does the perception of usefulness of different functions of an EAT change over time?

8.2.3 Emotional Competence as an Intervening Factor

As discussed in various parts of the thesis (see for instance Sections 1.1.3 and 3.5), some for of emotional expertise/competence/intelligence (Brackett et al., 2019; Hoemann et al., 2021; Scherer, 2007) may play a prominent role in determining whether an EAT can be instrumental in a computer-mediated learning environment. If, for instance, the definition and composition of emotional competence as derived from the Component Process Model (Scherer, 2007) is retained, it is possible to overlap the three sub-competences (appraisal, communication, and regulation) with specific functions of the Dynamic Emotion Wheel:

- The use of the sliders, corresponding to the cognitive evaluation of the situation, relies on the appraisal competence. An *adequate* assessment of the situation on relevant appraisal criteria can be helpful to initiate the meaning-making process.

- The expression of the subjective feeling through a lexicalized emotion or custom emotion term relies to the communication competence. The formulation of an accurate integration/categorization of the whole emotional experience is instrumental both from the internal, symbolic processing of emotional information (intra-personal meaning-making), but also from the inter-personal perspective of sending a strategic signal for others to understand.
- The emotion information available through the perceiving-monitoring side of the tool, at last, necessitates of an accurate reverse-engineering communicative strategy to infer the message behind the graphical representation, which can then be used for emotion regulation at the intra-personal or inter-personal levels.

Question 3 (Q3) therefore attempts at investigating whether some indicators of the use and/or perception of the EAT can be related to emotional competence. In this regard, congruently with the overall stance of the thesis that considers emotional competence something malleable, that can be acted upon, the construct will be measured through the use of a validated, performance-based test: the Geneva Emotion Competence (GECo) test (Schlegel & Mortillaro, 2018), illustrated in the Material part of the Methods section.

8.2.4 Recollection of Individual and Collective Emotional Experience

An assumption that may be subsumed from the importance of emotional awareness is that the emotional experience can have lasting effects on the meaning that is attributed to a (computer-mediated) learning experience (Brosch et al., 2013; Kensinger & Ford, 2020; Montagrin, Brosch, & Sander, 2013; Pool et al., 2015; Rimé, 2005). Question 4 (Q4) therefore aims at exploring whether the presence of an EAT contributes in *anchoring* the individual and collective affective experience during distance learning.

8.2.5 In Search of Emotion-As-Interaction Cues: Comparing Two Different Classes

Finally, a last topic of interest explored in the present contribution is in somehow transversal to the four previous questions. It concerns whether the use and perception of an EAT depends *mainly* on the individual characteristics of the students or are *also* determined by the interaction between students using the tool at the same time (Boehner et al., 2007; Dillenbourg et al., 2016). Interacting factors may relate, for instance, to learners not using the tool because nobody else does, or, on the contrary, because everybody else does. Another possible interacting factor may concern a general pattern in the emotional experience, which may result from a group-level form of influence (Barsade & Knight, 2015; Cheshin, Rafaeli, & Bos, 2011; E. R. Smith et al., 2007; Van Kleef et al., 2017). For instance, Cheshin et al. (2011) claim that emotional contagion can happen on virtual teams even if the communication between members is based only on text. One may therefore assume that the emotions of others can be all the more influencing on one's own emotional experience if there is a dedicated tool, overtly aimed at producing and perusing emotional information.

Taking advantage of the fact that two classes will use the EAT under very similar conditions – and assuming no systematic factor at play determining particular type of students in one class compared to the other (Pearl & Mackenzie, 2018a, 2018b; Rohrer, 2018) – wide differences between one class and the other would suggest the potential effect of interaction rather than individual characteristics alone. As previously stated, though, the exploratory nature of the study lacks the precision to quantify what is meant by *wide* differences. Consequently, results will be interpreted with respect to the available measures.

8.3 Methods

Following Simmons et al. (2012) suggestion to increase transparency in experimental contributions, I report how I determined the sample size, all data exclusions (if any), all manipulations, and all measures in the study.

8.3.1 Participants

33 students (22 women, 6 men, and 5 not discoled) of the course *Sciences et Technologies de l'Information et de la Communication I* (Fritz & Schneider, 2019) in the Master of Science in Learning and Teaching Technologies at Geneva University took part in the study (M_{age} = 32.96, SD_{age} = 7.78 with 7 not discolsed). Students belonged to two classes that took the one-semester course in two successive years during the period of the thesis: 16 students in the first class, and 17 students in the second, without overlapping. It is worth noting that a third cohort was originally planned to increase the sample size, but the study has not been implemented because this third cohort had to undergo an abrupt switch to an exclusive distance learning format due to the pandemic. Even though a completely online format would have been interesting for the purpose of the present study, the use of the EAT would have added up to an already complicated situation. Furthermore, the cohort would also have differed from the others in creating social links ,without the weeks with on-site sessions.

The use of the EAT during the course was warranted as a pedagogical activity, since the use of technological tools in learning situations is an integral part of the Master's program. Students had the choice, at the end of the course, either to sign a consent form allowing data to be used for research purposes, or to write on the same form only the ID they were given at the beginning of the study (without connection to their identity, see procedure below) so that all data associated with that ID could be erased. The overall sample size of the study is therefore determined by the number of students whose ID appears in at least one of the different sources of measure (see below) and has not been retracted via the *non-consent* form.

The population is clearly a convenience sample, but the choice has nevertheless both an explanation and a potential interest over and above the limitations. The explanation is preeminently of a technical nature. The EAT has never been adopted in a longitudinal study, where it must be seamlessly available even in the absence of the experimenter. Since the study implies comparison between two cohorts, it is therefore mandatory that technical problems should be signaled and repaired as soon as possible to maintain comparable settings. In this regard, MALTT students possess the technical know-how to identify and accurately describe malfunctioning in a web application, as well as a quick access to the technical team. The students are also exposed during the Master to an extensive use of different software and web applications, which helps to mitigate intervening factors related to technical skills.

The interest of the convenience sample is, ironically, its potential inconvenience. In fact, if the tool is adopted and considered as useful, results may be biased by the convenience sample, and therefore be taken with more than a grain of salt. On the other hand, if the tool is not adopted and considered of scarce usefulness, then the shortcomings are amplified, since even learners with an interest, habit, and technical know-how would not adopt it.

8.3.2 Material

Configuration of the Emotion Awareness Tool

The toolbox was configured to take into account the longitudinal use of the EAT, which may happen at any time during any day of remote learning, as well as with any number of students connected at the same time. The configuration of the toolbox, depicted in Figure 8.1, is thus implemented as follows.

For the expressing-displaying function of awareness, the same EATMINT circumplex was adopted. As a reminder, this affective space comprises 20 emotions organized over the two appraisal dimensions *Valence* and *Control/Power*. The rating of *Valence* was maintained as in the previous use of the circumplex, whereas for *Control/Power*, an attempt was made to reformulate the way it was prompted to increase its understanding and *independence* from the *Valence* dimension (see Section 4.2.3). The new wording adopted consisted in *Can you modify the situation?* Both dimensions were characterized by two opposite poles, labeled *Not at all* and *Yes, absolutely*. The choice and disposition of the subjective feelings is thoroughly depicted in Section 4.3.3.

For the perceiving-monitoring function of awareness, three word clouds were implemented. In a word cloud, words are depicted in a font whose size is proportional to the number of occurrences of that world, so that the more frequently used words appear bigger than the less frequently used ones. The perceiving-monitoring interface of the EAT comprised the following elements:

1. A *Self-Centered* word cloud, depicting the last 50 subjective feelings expressed by the participant herself;
2. A *Partners-Oriented* word cloud, depicting the last 100 subjective feelings expressed by the other members of the class, that is all the subjective feelings bar that of the participant herself;
3. A *Group-Oriented* word cloud, depicting all the subjective feelings expressed by the whole class since the first use of the EAT, that is, those of the participant herself, plus those of the other members.

The combination of the three word clouds sustain the two main social functions of emotions identified by Fischer and Mastead (2016): the affiliation and the distancing function (see Section 2.3). The top and center word clouds allow participant to compare her own emotions (top) with that of the class (center). In this way, the learner can position herself with respect to the other students (distancing). The bottom word cloud groups all the

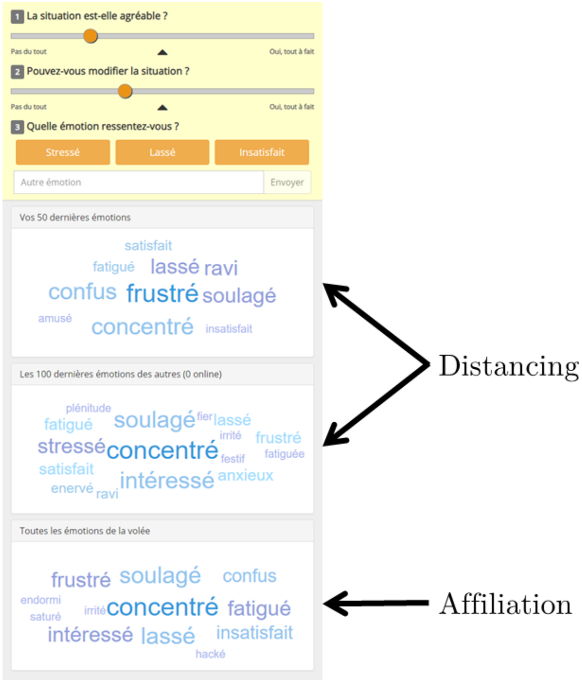


Figure 8.1: Interface of the EAT for a participant, depicting the expressing-displaying and perceiving-monitoring parts of the tool. For the perceiving-monitoring parts, the word clouds refer to the social functions of emotions identified by Fischer and Mastead (2016): affiliation and distancing (details in the text).

emotions of the class and is therefore more related to the affiliation function of emotions. In an attempt to increase the perception of others, the title of the central word cloud also reported how many other participants were currently online.

The number of emotions proposed by each word cloud is arbitrary, since there is no previous benchmark about frequency of use in general, and expression in particular. The word cloud also presents *a priori* shortcomings with respect to subjective feelings typed directly by participants, since for them to be grouped, they should be written in exactly the same way (*e.g.*, a small typo would isolate that expression). This is particularly relevant in French, since when emotion terms are used as adjectives, the declination changes according to the gender of the respondents (e.g. *intéressé* vs. *intéressée* with an addiontal letter *e* at the end). All things considering, though, word clouds are relatively known graphical representations, and convey an immediate and straightforward method of aggregation.

Compared with the interface used in the study illustrated in the previous chapter, thus, the differences concern only the perceiving-monitoring part of the EAT. First, all the emotional information was conveyed through the subjective feelings, with no trace of the cognitive evaluation. This choice is technically justified by the lack, at the time being, of a grouped representation of the appraisal dimensions, since the line charts used in the previous study are limited to individuals. Second, no temporal reference appeared at all, since the subjective feelings were categorized based on the frequency alone. This choice is justified by the fact that, in an asynchronous and non-collaborative context, the temporal reference conveys limited information, especially considering that there is no manifest link between the time and a specific task or situation involved. Third, except for the participant's own emotions, it was not possible to discern what specific colleague has expressed a particular feeling or cluster of feelings. Once again, this choice is technically imposed by the lack, at the time being, of a grouped representation of feelings, which is able to maintain agency without overcrowding the interface. The choice is nevertheless also a theoretical influence, since in such a setting, the emotions of the colleagues are *truly* at a group level (Barsade & Knight, 2015; E. R. Smith et al., 2007; Van Kleef & Fischer, 2015). More generally, without a previous benchmark about the number and frequency of emotions expressed in an asynchronous, individual situation, it was also difficult to establish grouping criterion (e.g. group by hour, by day, or by week). The grouped representation of emotions is an open issue that has just started to be investigate (see for instance Berset, 2018).

Emotion Awareness Usefulness Rating

The exploratory nature of the study was also considered the occasion to further advance the thesis overall perspective aiming at improving construct and internal validity in research gravitating around the concept of emotional awareness as defined in learning contexts. A central tenet of investigating a scientific construct is certainly the ability to measure it accurately, a topic that is receiving lately renewed interest (Boateng, Neilands, Frongillo, Melgar-Quiñonez, & Young, 2018; Flake & Fried, 2020). Validated scale that are close to the construct exists, but they are limited to some specific facets of emotional awareness. For instance, the Emotion Awareness Questionnaire (Rieffe et al., 2008) and the Levels of Emotional Awareness Scale (Lane, Quinlan, Schwartz, Walker, & Zeitlin, 1990), cited in Lavoué et al. (2020), relate to the individual context under clinical and personality psychology perspectives. They are not concerned with the inter-personal and instrumental use of emotional awareness in learning situations. There are also existing scales that measure the *social presence* construct, such as in Tu & McIsaac (2002) or Kreijns et al. (2011), but they do not specifically target the affective or emotional facet specifically. Finally, there are scales to measure emotion in academic contexts, such as in Pekrun et al. (2011) and Pekrun et al. (2016), but they aim at knowing the specific emotional experience of learners (e.g. through a list of discrete emotions), whereas in this case it is the meta-evaluation of emotional awareness that is at stake.

Based on these assumptions, the study introduces a tentative scale, the Emotion Awareness Usefulness (EAU), which aims at measuring emotional awareness specifically as a support for learning contexts. At the same time, the scale is intended to be as general as possible, without, for instance, making any reference to specific features of the EAT or the learning task. The aim, once again, is to provide an instrument that can foster comparison between studies. In this regard, the scale identify 7 dimensions that may be relevant to many situations in which emotional awareness is provided through an EAT:

1. **Frequency**. The frequency of use is a dimension used in different scales pertaining to Human-Computer Interaction and User Experience (MacKenzie, 2013; Tullis & Albert, 2013a), as it is the case in the System Usability Scale (Brooke, 1996). The

more frequently a tool is used, the more useful it is perceived, especially when the use is voluntary.

2. **Affordance**. Affordance in this context broadly refers to the actions available through the EAT and to what extent these actions are evoked by the tool (D. A. Norman, 2013; Rizzo, 2006; Suthers, 2006). For instance, the presence of an EAT may prompt users to express/share their emotions, something they would not do without the presence of the tool (e.g., Parkinson, 2008; Rimé, 2009; Van Kleef, 2018).

3. **Social Presence**. Social presence is a pivotal dimension in remote learning, providing support for learners isolation and feeling of loneliness (Jézégou, 2010; e.g., Kreijns et al., 2011). Perceiving the emotions of colleagues can help learners to remember there are others in the same condition as they are.

4. **Self-Understanding**. Having emotions a strong influence on intra-personal functions (e.g., Brosch et al., 2013; Levenson, 1999; Scherer, 2005), the presence of an EAT may contribute to a better self-assessment of the situation and its consequences on learners' behavior (Lavoué et al., 2020; Ruiz et al., 2016).

5. **Understanding Others**. The presence of the emotions of colleagues through the EAT may inform learners about what others are experiencing during the remote learning periods, providing information to build and update a mental model of the causes of and consequences on their behavior (e.g., Dillenbourg et al., 2016; Van Kleef, 2010, 2018; Van Kleef et al., 2017)

6. **Self-Other Comparison**. Comparing one's own emotions with that of the colleagues can provide useful information, especially in situation of incertitude (e.g., Eligio et al., 2012; Gaëlle Molinari, Chanel, et al., 2013; van de Ven, 2017). The presence of the EAT may facilitate and prompt this comparison.

7. **Self-Regulation**. Emotion regulation is a pivotal phenomenon that allows learners to modify their emotional experience using different strategies, such as suppression or reappraisal, in order to maintain instrumental emotional states and modify disruptive ones (e.g., Arguedas et al., 2016; J. J. Gross, 2014; J. J. Gross, 2015; Järvenoja et al., 2013). The presence of the EAT may facilitate regulatory processes both in terms of learners own emotions and that of the colleagues.

The items have been derived or adapted, when possible, from existing scales. Every item can be rated on a scale from 1 (strongly disagree) to 10 (strongly agree). The precise wording of each item is depicted in Table 8.1.

Table 8.1: Dimensions and respective items of the Emotion Awareness Usefulness (EAU) scale

#	Dimension	Item

1	Frequency	I used the tool frequently (e.g. every time I worked for the course).
2	Affordance	The use of the tool prompted me to share my emotions.
3	Social Presence	The use of the tool allowed me to feel less lonely during remote learning periods.
4	Self-Understanding	The use of the tool allowed me to better understand my emotions.
5	Understanding Others	The use of the tool allowed me to better understand the emotions of my colleagues.
6	Self-Other Comparison	The use fo the tool allowed me to compare my emotions with those of my colleagues.
7	Self-Regulation	The use of the tool allowed me to regulate my emotions.

The scale is very straightforward and with only a few items, specifically a single item per dimension. This is a potential shortcoming from a reliability and validity standpoint (Boateng et al., 2018), but, especially in a repeated measure design, brevity has been considered an important factor. Furthermore, the convenience sample of the study provides some leverage on the formulation of items. MALTT students are, for instance, familiar with terms such as *regulation*, which may be too technical for other populations and therefore would require a more explicit formulation.

System Usability Scale

The System Usability Scale (SUS) is a widely adopted scale that measures the usability of a tool (Brooke, 1996). It comprises 10 items, usually on a 5-point scale, but that can also be adapted to a 7-point range, which has been the case for this study. The 10 items of the scale are the following:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

The even items of the scale are reversed, so that a lower evaluation on the item corresponds to a greater perceived usability. The scale uses a system of coefficients that add up to obtain a score between 0 and 100. A more thorough discussion of the scale is available in the next chapter, where the usability score of the SUS will be compared between the usability test (Fritz, 2015) and the score obtained in the present study.

Geneva Emotional Competence Test

The Geneva Emotional Competence (GECo) test (Schlegel & Mortillaro, 2018) is a performance-based test that measures emotional competence as an ability, rather than a trait (Cherniss, 2010; Salovey & Mayer, 1990; Scherer, 2007). The test is primarily aimed at emotions in a workplace, but it may be considered that a working environment and a learning environment share many similarities in terms of intra-personal and inter-personal dynamics. The test is divided in 4 sub-competences parts, namely emotion recognition, emotion understanding, emotion regulation, and emotion management.

Emotion recognition determines the ability to infer the corresponding emotional state of a person using video clips of actors. Participants look and hear a professional actor expressing a pre-defined emotion using facial expression and pronouncing pseudo-words with a corresponding vocal prosody. Participants must identify the episode among several lexicalized emotions.

Emotion understanding determines the ability to infer the corresponding emotional state of a person based on a description of a situation. Participants read the details of an event that occurs to another person and must infer which emotion, among a list of discrete options, has been elicited by that event.

Emotion regulation determines the ability in engaging in adaptive (vs. disruptive) strategies to modify one's own emotional state. Participants read the description of a situation they must imagine has happened to them, which is meant to trigger disruptive emotional episodes. Respondents must then identify among four choices the two appropriate strategies, which allows the person to switch to a more adaptive emotional experience.

Emotion management determines the ability to adopt the better strategy to handle situations eliciting disruptive emotions in others. Participants read a vignette depicting a situation in which they interact with another person. The situation is meant to elicit in that other person a disruptive emotional response such as anger, irritation or misplaced happiness (*i.e.* the test considers that also *positively valenced* emotions may be disruptive). The participant can then choose between 5 different strategies to manage the emotional response of the other person, among which one is considered to be the more appropriate.

Each sub-test yields a score of accuracy. The 4 scores can then be combined to obtain an overall emotional competence score.

Individual and Class Perceived Frequency of Feelings

Considering the fact that the word clouds adopted in the perceiving-monitoring part of the EAT clearly define which feelings have been experienced more frequently than others, the study comprises a survey that asks participants, at the end of the semester, to rate the frequency by which (1) they recall having felt each of the 20 subjective feelings of the

EATMINT circumplex ; and (2) they recall their colleagues having felt each of the same 20 subjective feelings. In the survey, each subjective feeling is presented with a 5-point scale comprising *never*, *seldom*, *sometimes*, *often* and *very often*. Participants were also allowed to skip each particular feeling without rating the frequency both in the individual and the class surveys.

8.3.3 Procedure

All courses of the Master are organized in three periods per semester, which will defined by P1, P2 and P3. Each period is composed by a week of on-site courses, and 4-5 weeks of remote learning. In order to let students get familiar with distance learning, the use of the tool was integrated only from P2. In this way, students had P1 to get acquainted with the difficulties of distance learning, and could therefore better assess the usefulness of an awareness tool in general, and of an EAT in the specific case of this contribution.

During the on-site course of P2, students were informed that they would be asked to use the EAT as a corollary activity in the course. They were also informed that, beside the pedagogical interest of the activity, the use of the EAT was linked with a research topic and that data could therefore be used under specific conditions. The distinction between the participation to the *compulsory* pedagogical activity and the voluntary participation to the research was clearly explained, and students were asked to read a consent form that was linked into the private work-space of the course. They then drew an ID from a urn, which they would use for any interaction with the EAT or with the surveys. With the ID, which was unique for each student, they also received a common code for each class.

Expectancy Survey

At this point, each student was asked to fill the *Expectancy* survey in order to collect their perception of the use of an EAT before actually using it. The survey was simply introduced by this description:

> An emotion awareness tool is a tool that allows to share one's own emotions with other people in a computer-mediated context. The tool has two main function: (1) it allows the user to express her own emotions and make them visible to the other users who are using the tool; and (2) it allows the user to perceive the emotions of the other users who have access to the same tool. You will have access to an emotion awareness tool for the two remote periods of the course, so that you can use it while you are working on STIC I: while you are reading the pedagogical material, you are coding the devices for your exercises, or you are contributing to the Wiki.

The questions of the EAU were the same as described in the material above, except that they were transformed in a prospective tense. For instance, *I think I will use this tool frequently* or *I think the tool will prompt me to share my emotions*.

Demo Survey

Once filled the *Expectancy* survey, students were introduced to the EAT through a demo. They discovered the interface they will be using throughout distance periods of the course, and they could directly test the functioning of the tool by expressing emotions, knowing those will not be recorded. The general functioning of the tool (*i.e.,* the use of the appraisal dimensions and the choice of the subjective feeling) was explained. After 5 minutes of practicing with the tool, students took the *Demo* survey, which is exactly the same survey as the *Expectancy*, comprising the dimensions of the EAU in the prospective tense.

Set-Up of the EAT for Distance Periods

Towards the end of the on-site session in which the tool was presented, the set-up of the EAT for the actual use was organized. Since it was not possible to combine the EAT on the same interface with the various tools students use as part of the course, a generic web page was therefore the only flexible solution. Students saw how the window could be adapted and put beside another window (*e.g.,* of a software or of another web page) in order to have the EAT close to the task at hand. To ease the access to the EAT, students created a bookmark in their browser that would automatically log them in with their unique ID and the code of the class.

Practically, then, students were supposed to open their browser, click on the bookmark pointing to the EAT, resize the window and place it beside the activity they were performing for the course. Or, alternatively, keep it minimized on their operating system task area and maximize it on recall. In both cases, the use of the EAT required a deliberate action outside the *normal* work-flow of the course.

Halfway Survey

During the on-site course of P3, students filled in the *Halfway* survey, after a whole period (i.e. 5 weeks) of use of the tool during remote learning. The survey comprised the EAU dimensions in retrospective tense (*e.g., I have used the tool frequently* or *The tool has prompted me to share my emotions*), as well as an open-ended question in which students could provide additional information about their experience with the EAT.

Final Survey

Students filled the *final* survey during the first on-site course of a follow-up course (STIC II) in the next semester, that is after 9 weeks from the *halfway* survey. The long period is the result of 5 weeks of *normal* remote learning, interrupted by 2 weeks of Christmas' leave in December (in which students often works, though), and 2 weeks of end-of-semester leave. The EAT was available until the formal end of semester. It has been decided to ask students to fill in the *final* survey on-site, even with two weeks delay compared to the end of the semester, in order to maximize data collection. The *final* survey comprised:

- The EAU survey in retrospective tense (same as *halfway* survey);

- The System Usability Scale (Brooke, 1996), but with a 7-point scale rather than the usual 5-point scale;
- A survey asking participants to rate the frequency with which they have experienced the 20 subjective feelings belonging to the underlying affective space of the DEW;
- A survey asking participants to rate the same feelings, but with respect to the frequency with which their colleagues have felt them during the remote periods;
- An open-ended question in which students could provide additional information about their experience with the tool.

Reminders

Reminders to use the EAT during remote learning periods were dispatched twice per periods (P2 and P3) within messages in the private space of the course. The reminders were integrated into wider communications, for instance the feedback of an exercise.

Geneva Emotional Competence Online Survey

In the private work-space of the course, students could find a link to the Geneva Emotional Competence (GECo, Schlegel & Mortillaro (2018)) test. The presence of the link was reminded at each on-site course, but students were clearly informed that the test was exclusively part of the research, so they were not forced to take it. Students could therefore take the test anytime during the P2 or P3 periods. Considering that there is no evidence yet that the use of an EAT in this context could improve the emotional competence of a person, especially in a performance-based test, this option left more time for students to take a long test during a very active period of learning. Students deciding to take the test had only to provide their anonymous ID.

8.3.4 Exclusion Criteria

Having no previous reference for data collection, *a priori* exclusion criteria were difficult to formalize. Since there was the possibility of students dropping out either from the Master or from the research activity (e.g. refusing to fill-in the surveys), participants not having filled both the *halfway* and the *final* survey will be considered as if they dropped out and will thus be excluded from data analysis.

8.4 Results

Results are based on $N = 30$ participants. 3 students did not fill neither the *halfway* nor the *final* survey and were therefore excluded from the analysis. As mentioned above, it was considered that they have dropped off either from the course or the research. Table 8.2 depicts the number of participants retained for each class across the 4 longitudinal surveys *expectancy*, *demo*, *halfway* and *final*, for which the two classes have very similar – if not equal – sample sizes.

Table 8.2: Number of participants retained for each class in total, and with respect to the 4 longitudinal surveys.

	Class 1	Class 2
Number of students retained per class	15	15
Number of students filling the expectancy survey	15	15
Number of students filling the demo survey	15	14
Number of students filling the halfway survey	13	15
Number of students filling the final survey	13	13

8.4.1 Measures About the Use of the EAT

Data about the adoption and use of the tool are limited here to the number of emotions expressed. This is clearly a shortcoming, since, potentially, participants could have perused the emotional information available on the interface without expressing any emotion (ever, or only during a specific session). As illustrated in the presentation of the toolbox, it would have been possible to use the accesses of participants to provide a more accurate estimation of the adoption and use. Nevertheless, the feature was developed too late, once the design of the study was already finalized, and in the information to participants, as well as in the consent form, there was no mention of recording participants accesses. For this reason, data related to accesses have been erased.

Overall, participants expressed 374 emotions through the EAT, that is a mean of $M = 12.47$, 95% CI $[6.05, 18.88]$ ($SD = 17.18$). One student in Class 1 and three students in Class 2 did not express any emotions at all. As mentioned above, this does not rule out the possibility that they nevertheless have accessed the tool. They are therefore kept into the analysis.

In comparing the two classes, the number of emotions expressed is similar with respect to the central tendency ($M_{class1} = 11.47$ against $M_{class2} = 13.47$), but differs greatly in variation ($SD_{class1} = 21.69$ against $SD_{class2} = 11.79$). The greater SD of Class 1 results in particular by a single participant that expressed 86 emotions. In Class 2, the greatest number of emotions expressed is 34. Taking the median as a more robust reference of comparison, the difference between Class 1 ($Mdn = 3$) and Class 2 ($Mdn = 12$) is much more evident. Figure 8.2 compares the expression of emotions between the two classes.

The use of the tool with respect to the longitudinal duration of the course is depicted in Figure 8.3. The classes have similar profiles in the cumulative number of emotions expressed over time, alternating phases in which they express emotions with periods of pause.

Overall, it is safe to assume that Class 2 made a more thorough and homogeneous use of the EAT in expressing emotions compared to Class 1, even though the use in expressing emotions remains, overall, limited to a few participants that expressed most of the emotions.

8.4.2 Emotion Awareness Usefulness

Being the first adoption of the Emotion Awareness Usefulness (EAU) scale, it is useful to start by evaluating the sensibility of the scale in terms of overall dispersion for each of the 7

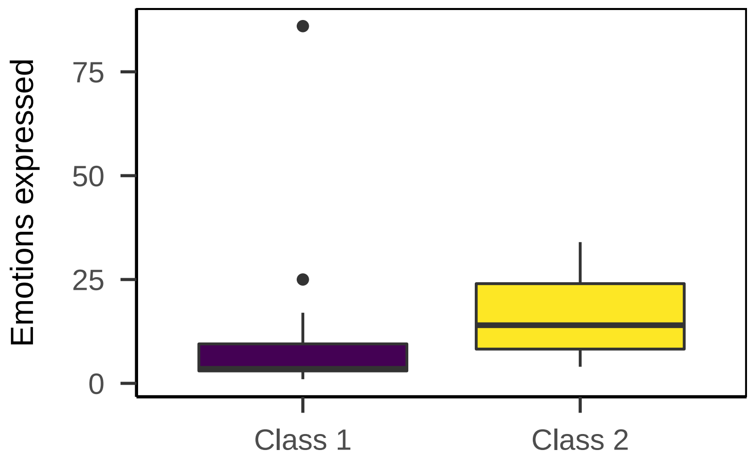


Figure 8.2: Boxplot comparing the expression of emotions between the two classes.

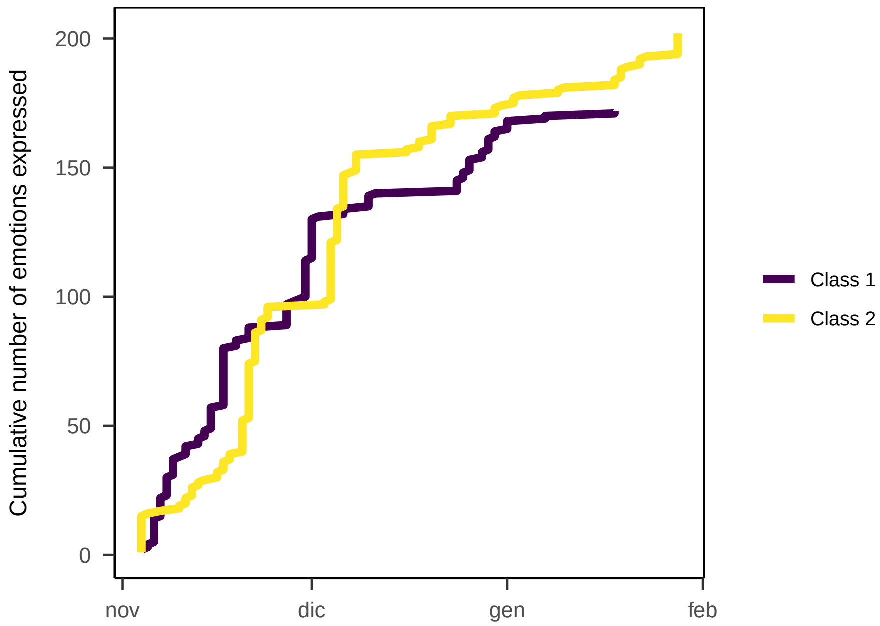


Figure 8.3: Cumulative number of emotions expressed through the EAT

underlying dimensions, namely *Frequency*, *Affordance*, *Social Presence*, *Self-Understanding*, *Understanding Others*, *Self-Other Comparison*, and *Self-Regulation*. In Figure 8.4 every circle represents one of the $N = 798$ ratings, which was made on a scale from 1 to 10, for each dimension across surveys and participants. The dispersion for each of the 7 dimensions spans the entire range of the rating, suggesting that each dimension has a good sensibility.

This preliminary assessment therefore corroborates the interest to investigate more specific patterns in the rating of each of the 7 dimensions with respect to the change over time and differences between classes. Table 8.3 depicts means and standard deviations of the 7 dimensions stratified for the two classes across the 4 longitudinal surveys.

Figure 8.4: Overall ratings of each of the 7 EAU dimensions, $N = 798$ ratings. Bars represent 95% confidence intervals around the overall mean.

Table 8.3: Means and (standard deviations) of the EAU ratings across longitudinal surveys.

group	Expectancy	Demo	Halfway	Final
Frequency				
Class 1	6.47 (2.59)	6.33 (2.69)	5.64 (2.27)	3.31 (2.10)
Class 2	4.67 (1.72)	4.93 (2.40)	3.87 (2.77)	3.92 (2.90)
Total	**5.57 (2.34)**	**5.66 (2.61)**	**4.72 (2.66)**	**3.62 (2.50)**
Affordance				
Class 1	7.20 (1.47)	7.13 (1.88)	5.86 (2.35)	4.46 (1.45)
Class 2	5.93 (2.37)	6.00 (2.63)	4.60 (3.07)	4.23 (2.77)
Total	**6.57 (2.05)**	**6.59 (2.31)**	**5.21 (2.77)**	**4.35 (2.17)**
Social Presence				
Class 1	5.07 (2.63)	5.40 (2.56)	3.14 (2.28)	2.15 (1.07)
Class 2	6.60 (2.35)	6.21 (3.26)	3.13 (2.29)	3.46 (2.44)
Total	**5.83 (2.57)**	**5.79 (2.90)**	**3.14 (2.25)**	**2.81 (1.96)**
Self-Understanding				
Class 1	5.00 (2.27)	4.40 (2.56)	4.43 (2.68)	3.69 (2.06)
Class 2	4.13 (2.17)	4.43 (2.41)	3.00 (1.89)	3.31 (2.32)
Total	**4.57 (2.22)**	**4.41 (2.44)**	**3.69 (2.38)**	**3.50 (2.16)**
Understanding Others				
Class 1	7.00 (2.24)	6.87 (2.03)	4.50 (2.77)	4.69 (2.18)
Class 2	6.87 (2.13)	7.00 (1.84)	3.80 (2.73)	3.62 (2.14)
Total	**6.93 (2.15)**	**6.93 (1.91)**	**4.14 (2.72)**	**4.15 (2.19)**
Self-Other Comparison				
Class 1	7.87 (1.81)	8.00 (1.46)	4.93 (2.53)	4.77 (2.59)
Class 2	7.53 (2.47)	7.79 (2.39)	5.07 (3.08)	4.46 (2.63)
Total	**7.70 (2.14)**	**7.90 (1.93)**	**5.00 (2.78)**	**4.62 (2.56)**
Self-Regulation				
Class 1	4.00 (2.67)	4.40 (2.56)	2.79 (2.12)	2.54 (1.39)
Class 2	4.47 (2.45)	4.29 (2.02)	2.53 (1.73)	3.31 (2.25)

In order to assess the presence of effects in the above ratings, a linear mixed-model, also known as multilevel linear models (Finch, Bolin, & Kelley, 2019; West et al., 2015), was fitted to the data with the following parameters. The dependent variable is the singular value, from 1 to 10, expressed on each dimension of the EAU in the 4 surveys. The fixed covariates are (1) the longitudinal survey with 4 levels (*Expectancy*, *Demo*, *Halfway*, and *Final*); (2) the specific dimension of the EAU with 7 levels (*Frequency*, *Affordance*, *Social Presence*, *Self-Understanding*, *Understanding Others*, *Self-Other Comparison*, and *Self-Regulation*); and (3) the group each student belongs to with 2 conditions (*Class 1* or *Class 2*). All the two-way interactions between pairwise covariates, as well as the three-way interaction, were also fitted in the model. The random covariates account for the nested structure of the observations: the repeated measure of the participant is nested inside the class to which the participant belongs.

The use of a linear-mixed model is warranted by a better flexibility compared to repeated-measure ANOVA, especially in case of missing data. Furthermore, linear-mixed models allow to account both for the repeated measures per participant and the nested, hierarchical structure of residuals, with participants potentially influenced by the use of the tool made by the colleagues in their class, which would violate the non-independence of residuals in ordinary least square regression. Finally, it allows to keep each dimension separate rather than averaging over a single score, which will loose an interesting source of variance with respect to the 7 dimensions of the EAU (see Finch et al. (2019); McElreath (2020); Singmann & Kellen (2020); West et al. (2015); Brown (2021) for a more comprehensive overview of linear mixed models compared to ordinary least square regression).

The linear mixed model analyzing the score on the EAU scale was fitted using the mixed function of the Afex (Singmann et al., 2020; Singmann & Kellen, 2020) R package version 1.1.1. A Type III analysis of variance of the multilevel linear model detected effects for the longitudinal survey and the dimension of the EAU scale, but not for the group. Two-way interactions were detected between the group and the EAU dimension as well as between the EAU dimension and the longitudinal survey, but not between the group and the longitudinal survey. Finally, the three-way interaction between group, longitudinal survey and EAU dimension was not detected. Results are depicted in Table 8.4 using Kenward-Roger approximation for computing the p-value (Luke, 2017).

Table 8.4: Results of a Type III ANOVA on the fitted multilevel linear model using Kenward-Roger approximation for computing the p-value

	num Df	den Df	F	Pr(>F)
group	1	27.986	0.557	.462
survey	3	716.753	63.495	<.001
dimension	6	713.993	29.403	<.001
group:survey	3	716.753	1.173	.319
group:dimension	6	713.993	3.769	.001
survey:dimension	18	713.993	2.085	.005
group:survey:dimension	18	713.993	0.664	.848

Results are graphically depicted in Figure 8.5, in which the EAU evaluation is stratified by the 7 dimensions of the scale and the 4 longitudinal surveys, as well as divided between the 2 classes. Data show high *expectancy* ratings for the *Frequency*, *Affordance*, *Social Presence*, *Understanding Others*, and *Self-Other Comparison* dimensions, but not for *Self-Understanding* and *Self-Regulation*. The ratings are generally maintained even after the *demo* surveys and tend to decrease with the actual use of the tool in the *halfway* survey, to remain then more or less stable even in the *final* survey. Data also show that the two classes basically overlap on all dimensions and across longitudinal surveys. The group per dimension interaction is probably yielded by the *Frequency* and *Affordance* dimensions, for which the two classes differ the most, but it does not seem to play an important role in differentiating EAU ratings. As a consequence, the factor related to the class is dropped in post-hoc contrasts, which will focus only on the difference between the *Final* and the *Expectancy* surveys.

Results of the post-hoc contrasts averaged over group and stratified across EAU dimensions are illustrated in Table 8.5. Contrasts detect differences between the *Final* and the *Expectancy* surveys in every dimension except the *Self-Understanding* one. All detectable differences highlight a decrease in perceived EAU, with the greatest decrease for *Self-Other Comparison* and *Social Presence* (almost 3 rating points), followed by *Understanding Others* and *Affordance* (more than 2 rating points), and finishing with *Frequency* and *Self-Regulation* (more than 1 rating point).

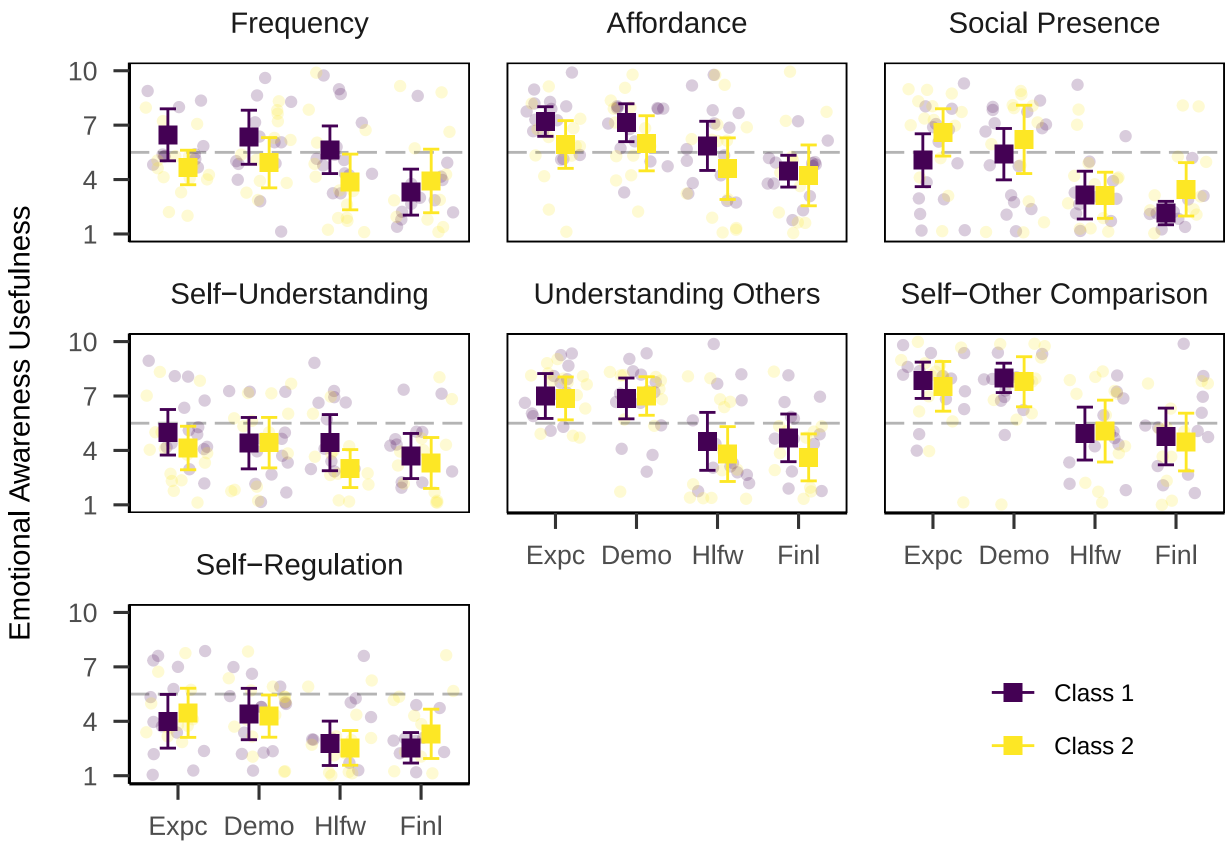


Figure 8.5: EAU rating over longitudinal surveys stratified by dimensions and grouped by class. Bars represents 95% confidence interval and the dashed gray line the median point of the scale.

Table 8.5: Contrasts between the Final and the Expectancy surveys for each of the 7 dimensions of the EAU scale averaged over the two classes.

contrast	dimension	estimate	SE	df	z.ratio	p.value
Final - Expectancy	Frequency	-1.754	0.519	Inf	-3.381	<.001
Final - Expectancy	Affordance	-2.024	0.519	Inf	-3.899	<.001
Final - Expectancy	Social Presence	-2.829	0.519	Inf	-5.451	<.001
Final - Expectancy	Self-Understanding	-0.870	0.519	Inf	-1.676	.094
Final - Expectancy	Understanding Others	-2.583	0.519	Inf	-4.976	<.001
Final - Expectancy	Self-Other Comparison	-2.888	0.519	Inf	-5.564	<.001
Final - Expectancy	Self-Regulation	-1.113	0.519	Inf	-2.145	.032

8.4.3 Structure and Reliability of the Emotion Usefulness Scale

Even though the use of multilevel linear model is more adequate for a fine grained assessment of the tentative scale, it is worth exploring whether the Emotion Awareness Usefulness scale

(EAU) proposes structural features that may be of interest for future use. In this regard, the way the scale has been administered is somehow atypical. On the one hand, participants rated the scale on multiple occasion, which is consistent with the test/re-test paradigm. On the other hand, the conditions in which the scale has been administered where obviously not the same, since the aim of the contribution was to measure the change of perception over time rather then the consistency of the scale. For an exploratory purpose, it could be informative to relax some of the habitual boundaries in reliability measures and *pretend* that each administration of the survey was unique, even when the rater was the same. This choice is warranted by two elements. First, it increases the sample size of measures compared to taking just one of the four survey in which the scale has been administered. Second, if the repeated measure of the administration is taken into consideration, there will be essential overlapping with the multilevel analysis performed above. For these reasons, all $N = 114$ administrations of the scale will be considered in the analysis.

Two types of analyses will be conducted to explore the structure and reliability of the EAU. First, it is worth assessing to what extent the scale can be considered as uni-dimensional, that is, measuring a single underlying latent variable, represented in this case by the perceived usefulness of disposing of emotional awareness. Second, it is also worth exploring whether there is an underlying structure of more than one factor.

Uni-Dimensionality

For determining the existence of a single, common factor, there is a growing consensus in considering that Cronbach's α is not the most adequate choice (see for exemple Hayes & Coutts, 2020; Revelle & Condon, 2019; Sijtsma, 2009). In this regard, Revelle & Condon (2019) suggest that more than one indicator of the psychometric properties of the scale should be reported. The table below illustrates indicators of reliability computed using the psych R package (Revelle, 2021), used here in version 2.2.5.

Table 8.6: Reliability measures of the EAU scale

ω_h	α	ω_{tot}	Uni	r.fit	fa.fit	max.split	min.split	mean.r	med.r
0.6	0.86	0.9	0.86	0.92	0.93	0.88	0.71	0.46	0.35

All indicators seem to converge on moderate to high values, which corroborates the reliability of the scale. For the purpose of this study, particular attention is focused on the *Uni* indicator, which is computed through the *unidim* function of the psych package (*ibid.*). The function is described as follows:

> There are a variety of ways of assessing whether a set of items measures one latent trait. unidim is just one more way. If a one factor model holds in the data, then the factor analytic decomposition F implies that FF' should reproduce the correlations with communalities along the diagonal. In this case, the fit FF' should be identical to the correlation matrix minus the uniquenesses. unidim is just the ratio of these two estimates. The higher it is, the more the evidence for unidimensionality.

Results of fitting the *unidim* approach to the EAU yields a unidimensionality index of 0.86, with a fit of the average correlation to the matrix of 0.92. This may suggest that the scale can, in principle, be reduced to a unidimensional latent factor.

Exploratory Factor Analysis

Even though the scale provides a good approximation for a single factor, it is still worth examining if more than one factor better fit the structure. The exploratory factor analysis will be conducted in two steps using again the psych R package (Revelle, 2021) version 2.2.5. First, a scree test is conducted to determine the number of factors that account for the greatest variance. Then, this number of factors is used in an exploratory factor analysis to retrieve the loading of the seven dimensions of the EAU on the suggested number of factors.

The scree test pictured in Figure 8.6 suggests that three factors account for a fair amount of variance, with small gain achieved by adding a fourth or a fifth dimension.

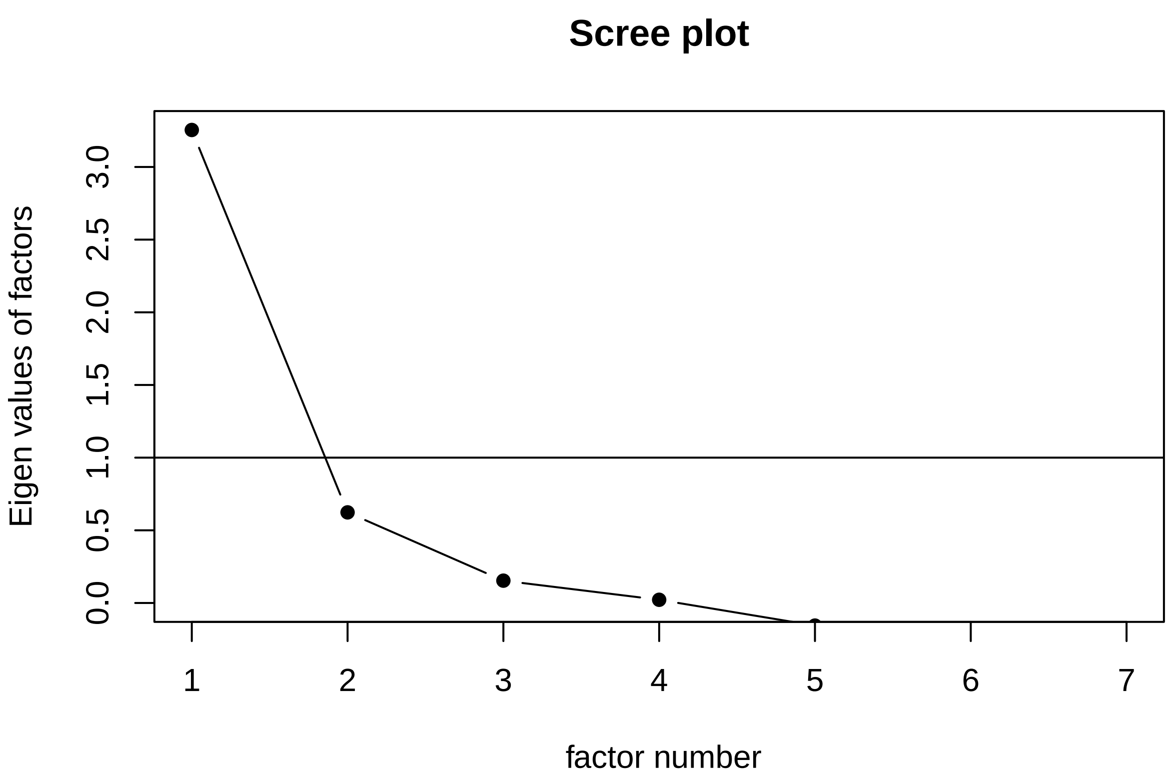


Figure 8.6: Scree plot based on all the EAU surveys administered. $N = 114$

An exploratory factor analysis with three factors, using the *minimum residuals* method of extraction with the *oblimin* rotation, results in the following loading of dimensions:

1. The first factor comprises *Social Presence*, *Understanding Others*, and *Self-Other Comparison* dimensions, that is the more *partner(s)-oriented* dimensions.

2. The second factor includes *Self-Understanding* and *Self-Regulation* dimensions, that is the more *self-centered* dimensions of emotional awareness. The second factor also relates to *Social Presence*, which also loads on the first factor, and *Frequency*, shared with the third factor, but both with a lower coefficient;
3. The third factor represents *Frequency* and *Affordance* dimensions, that is the more *usability related* dimensions.

The diagram in Figure 8.7 reports the coefficients of loading. The overall reliability scores are Cronbach's $\alpha = 0.86$, $\omega_h = 0.66$, and $\omega_t = 0.93$, which suggest good reliability of the scale. Incomplete and tentative, the EAU seems nevertheless a rather promising scale that can be expanded in further research.

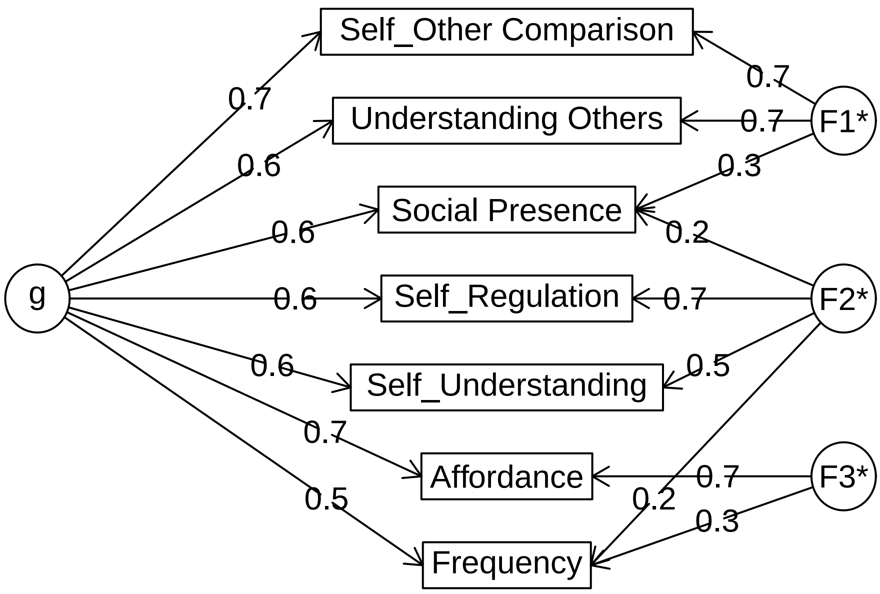


Figure 8.7: Graphical representation of the exploratory factor analysis

8.4.4 Emotion Awareness Usefulness and Emotional Competence

The study also included the possibility to take the Geneva Emotion Competence (GECo) test (Schlegel & Mortillaro, 2018) anytime before the end of the semester. As a reminder, the GECo is a performance-based test measuring participants' emotional competence on four sub-competences: emotion recognition, emotion understanding, emotion regulation, and emotion management. Given the test was a corollary activity related only to the research – and that the test is demanding in terms of time and effort due to its accuracy – only $N = 11$ participants took it. This limits the possibility of exploring relationship between emotional competence and other indicators collected throughout the study beyond a descriptive perspective. In the meantime, participants that did take the test used their precious time during an intense phase of their education, which deserves gratitude and consideration.

Given the overall limited use that has been done of the EAT, the more interesting measure at hand to be related to emotional competence is the prospected usefulness measured in the *Expectancy* survey using the EAU scale. The unidimensionality coefficients exposed above suggest that the average score of the EAU can be, at least in the exploratory context, retained as an indicator of perceived usefulness of emotional awareness and therefore represent the measure of interest. With these premises, the most informative test retained in this case is the use of a Bayesian linear model using informed priors (Dienes, 2016; Etz & Vandekerckhove, 2018; Makowski, Ben-Shachar, Chen, & Lüdecke, 2019; McElreath, 2020; Wagenmakers et al., 2018). Bayesian analysis is gaining interest in the social sciences (*ibid.*) and in this specific case, where the analysis is exploratory, it may represent a good attempt to obtain the most information from the small sample at hand.

In the model, thus, the outcome variable is represented by the unidimensional average of the EAU score and the predictors are the four sub-competences of the GECo test. The aim of the test is to explore whether specific sub-competences seem to be more related to the overall expectation of the perceived potential of emotional awareness. As informed priors, the model uses the means and standard deviations observed in Study 1 in Schlegel & Mortillaro (2018), where the GECo is validated. The measures are reported in Table 1 of the original article on page 9. Among the 5 studies on which the GECo has been validated, the first is the one whose population is closer to the participants in this study. Participants to the GECo validation's Study 1 were in fact 149 students from different faculties at the University of Geneva, and the test was administered in French as it was the case for this study. In Study 1 of the GECo validation, the score of the sub-competences were the following: $M_{Recognition} = 0.66$ ($SD = 0.13$); $M_{Understanding} = 0.70$ ($SD = 0.16$); $M_{Regulation} = 1.12$ ($SD = 0.22$); $M_{Management} = 0.45$ ($SD = 0.18$). For the $N = 11$ participants in this study, the scores to the same sub-competences were: $M_{Recognition} = 0.73$ ($SD = 0.11$); $M_{Understanding} = 0.78$ ($SD = 0.17$); $M_{Regulation} = 0.57$ ($SD = 0.10$); $M_{Management} = 0.55$ ($SD = 0.15$). The regulation sub-competence is therefore the only one that seems to deviate considerably from the validation's sample, with an observed mean that is half the score observed in Schlegel & Mortillaro (2018). Notwithstanding this huge difference – for which, honestly, an explication is difficult to provide – the priors seem adequate to normalize the parameters and provide an informed framework for a Bayesian linear model (McElreath, 2020).

The analysis was performed using the rstanarm version 2.21.3 R package (Goodrich, Gabry, Ali, & Brilleman, 2022). The model was estimated using the Markov Chain Monte-Carlo sampling, with 4 chains of 10'000 iterations and a warm-up of 5'000. The prior were scaled internally by the model to provide some leverage considering the small sample size at hand. Otherwise, the poster distribution would have basically confirmed the strong priors

(Goodrich et al., 2022). The adjusted standard deviation for the prior was respectively of $SD_{Recognition} = 1.47$, $SD_{Understanding} = 1.16$, $SD_{Regulation} = 2.76$, and $SD_{Management} = 1.48$. Table 8.7 reports the summary of the posterior distribution. For each parameter, the Probability of Direction (pd), the Region of Practical Equivalence (ROPE), and the Bayes Factor (BF) are reported (Makowski et al., 2019). Contrary to frequentist's statistics where significance is commonly determined with the p-value, in Bayesian analysis there are more options and, for the time being, not established conventions to interpret results with specified benchmarks (Makowski et al., 2019).

Table 8.7: Results of a Bayesian linear model analysis with the standardized average of the EAU scale as outcome variable, and the four GECo sub-competences as parameters.

Parameter	Median	CI	pd	ROPE	BF	Rhat	ESS
Intercept	4.35	[0.58, 8.10]	0.99	0.00%	3.86	1	5621.3
Recognition	-0.13	[-2.86, 2.55]	0.54	0.07%	0.86	1	4010.2
Understanding	0.68	[-1.37, 2.60]	0.74	0.08%	0.92	1	4660.5
Regulation	1.18	[-3.22, 5.73]	0.70	0.04%	0.85	1	5372.8
Management	1.63	[-0.99, 4.26]	0.89	0.03%	2.03	1	3365.4

Even without pre-defined benchmarks, though, as one may expect from the small sample size at hand, all the parameters remain uncertain and the test provide inconclusive results about the possibility that the perceived emotional awareness usefulness can be related to emotional sub-competences. It is nevertheless interesting to note that, assuming the parameters would hold in a larger sample, emotion recognition has virtually no effect, whereas regulation and management have effects that are respectively almost twofold, and over twofold that of understanding. As a reminder, emotion recognition was performed through efferent manifestations that would normally be available in face-to-face interactions, and its lack of effect could be considered as supporting the need for an EAT to represent emotion with a structure rather than efferent cues. The importance of regulatory processes would also be corroborated, with an interesting perspective coming from the greater effect of inter-personal emotion regulation compared to intra-personal emotion regulation. These are nevertheless only wild speculations for the time being. The only relevant element to retain may therefore be that the GECo test and its sub-competences are viable measures to further investigate this kind of effects.

8.4.5 Perceived Usability of the EAT

$N = 26$ participants filled the System Usability Scale (SUS) survey (Brooke, 1996) at the end of the semester. The observed score has been of $M = 72.82$, 95% CI $[68.02, 77.62]$ ($SD = 11.89$), which is considered *Good* according to the stratification proposed by Bangor, Kortum and Miller (2009).

Figure 8.8 illustrates the score obtained on each of the 10 items of the SUS for the two classes. The score takes into account the reverse order for even items and is pondered on the 7-point scale used in the present study. As in the case of the EAU, ratings on the SUS are also consistent between classes. More specifically, data clearly show that the first item, inherent to the frequency of use, is far below the other items, which have received overall a high rating. Items 5 (integration of different functions of the system) and 9 (confidence in using the system) also have received lower ratings, even if not as low as the first item.

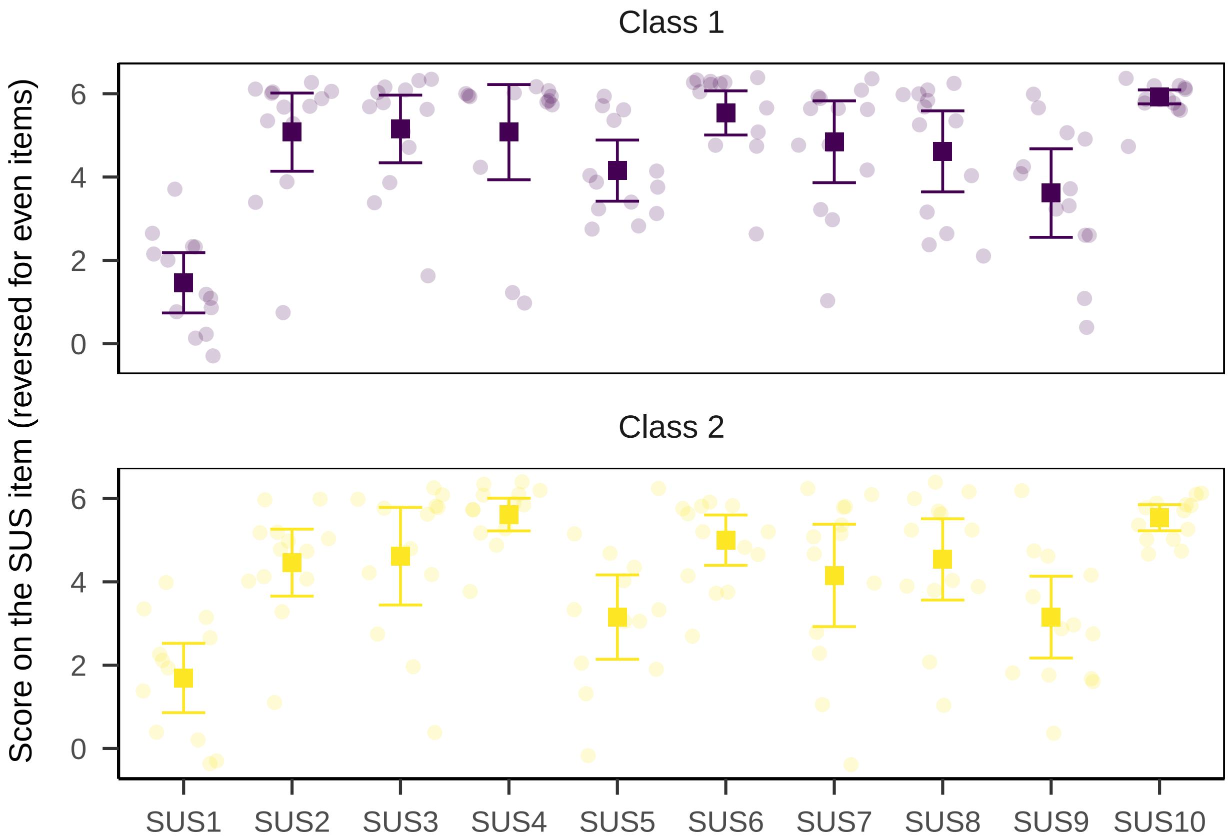


Figure 8.8: Rating of the individual items of the System Usability Scale (SUS) for N = 26 participants.

A more detailed analysis on perceived usability is presented in the next chapter, which will compare the score with the usability test (Fritz, 2015). More accurate benchmarks will also be adopted to assess the individual items of the scale (Lewis & Sauro, 2018).

8.4.6 Recollection of Individual and Collective Emotions

In the *final* survey, that is approximately two weeks after the end of the semester – the day the instance of the toolbox was closed – participants were asked to recall the frequency with which (1) they have experienced the 20 subjective feelings of the underlying EATMINT affective space, and (2) their class as a whole has experienced the same 20 subjective feelings. The sparse use in expressing emotions limits the interest in assessing whether there is a recollection of the individual and collective emotions shared with the EAT and is therefore reported here primarily to comply with the disclosure of all the measure collected. Figure

8.9 compares the frequency with which the participant herself thinks she has experienced each emotion (Self) with the frequency that she attributes to the whole group (Class). The gray line available on each comparison represents the *observed* frequency with which the subjective feelings has been expressed through the EAT. This frequency is computed by mapping the expressed feelings that have been chosen the most (e.g. *Attentive* and *Interested*) as the *Very often* frequency, whereas the least expressed feelings (e.g. *Surprised* and *Disgusted*) represent the lower bound, and correspond to the *Never* frequency. *Envious* and *Bored* do not have a horizontal line because there were not even a single expressed emotion with these subjective feelings.

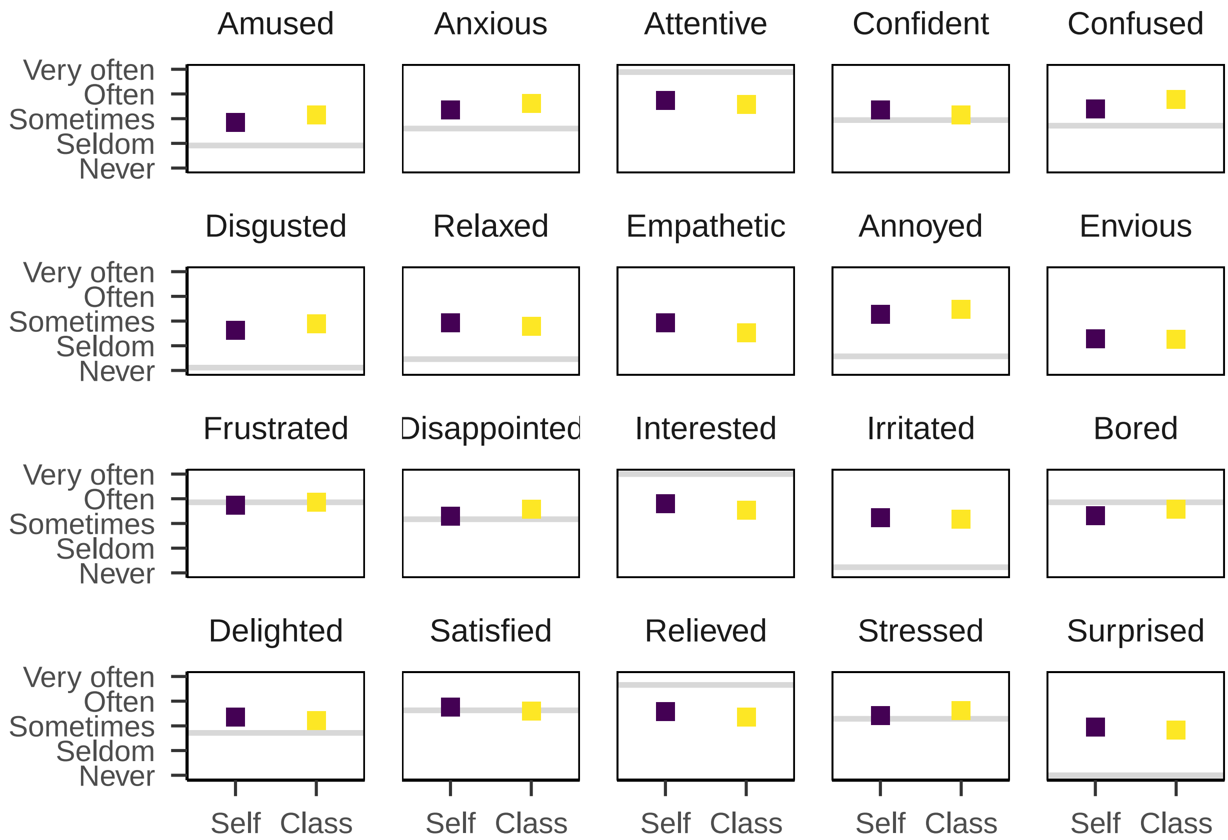


Figure 8.9: Comparison between the frequency with which each subjective feeling has been recollected as being perceived by the learner herself (Self) or attributed to the whole group (Class). The gray line represents the *observed* frequency mapped on the number of expressions through the EAT.

Data basically show that participants tended to rate a similar frequency for their colleagues based on their own frequency, a phenomenon already encountered in other contributions (Eligio et al., 2012; Krueger & Clement, 1994; Toma et al., 2010). When participants lack the knowledge of their partners emotions, they use themselves as the *measuring stick*. The frequency is also mostly independent from the *observed* frequency, except for *Frustrated*, *Bored*, and *Satisfied*, for which the *observed* and the *class* estimation overlap. The estimation for the group and the *observed* frequency may, in principle at least, coincide, since the graphical representation of the whole class – the word cloud – was overtly aimed at the relative occurrence of subjective feelings. So, if the word cloud about the class as a whole was retained in memory, at least this measure could potentially be precise. The fact that this is mostly not the case seems to suggest that, simply, there has been no recollection of the emotional experience from the EAT, and estimation are based on individual background knowledge or some form of retrospective theory of mind (Hall et al., 2018; Wellman, 2018).

8.4.7 Open-Ended Question

In the *halfway* and *final* surveys, participants had the possibilities to add general commentaries to the overall questionnaire, with a free text field corresponding to the overarching question *Would you like to add any remark or comment on the study as a whole?* The open-ended question received 21 answers in the *halfway* survey, and only 10 in the *final* survey. Even though the overall approach to the study is nomothetic rather than idiographic, the fact that the tool's adoption has been sparse overall suggests that these questions could provide some useful reasons.

In this regard, most of the answers both in the *halfway* and *final* surveys point out that students forgot to use the tool, primarily because they were already overwhelmed by the learning task. A few students, though, precised that the lack of use was also influenced by the presence of alternative communication channels used by the class, as it is clearly stated in this answer:

> At first, I forgot to use it and realized it late in the period. Afterwards, I didn't do it, because I use [the group's messaging platform] Slack to exchange and this tool is sufficient for me to manage my emotions. Using another tool would have generated an additional mobilization of resources and [would] therefore [have been] counterproductive for me.
> — Verbatim of the answer, translated from original French.

Other students highlighted the fact that the moment-to-moment use of the tool interfered with their learning activity, diverting their attention. Even if they did manifest their emotion, those were rather prospective or retrospective (Pekrun, 2006; Pekrun & Perry, 2014). Furthermore, the sparse use of the tool even produced the very opposite effect it was intended to obtain, that is, increase social presence. As stated by a student:

> I used the tool and I like the concept but I only found myself online once with my other colleagues which actually made me feel quite alone in my sharing moments. This is the opposite of what I expected because I thought I would feel more connected to my colleagues and my emotions in the moment.
> — Verbatim of the answer, translated from original French.

A few students nevertheless provide positive feedbacks about the presence of the tool, even if they regret the sparse use made of it by themselves or their colleagues. The feedback that more closely adheres to the tool overall purpose, though, provide an attentive analysis of its use in terms of the role the appraisal criteria played in its adoption:

> It was a great tool to use. I particularly appreciated to have this tool in order to force myself to analyze my emotions while working on [the course] topics. I particularly appreciated the question about the capability to change the situation, which was a good opportunity to take a step back and understand what factors led to the situation and how I could influence them. It was also particularly interesting to be able to see the last emotional states and the most used. I would be happy to continue the survey in the future period.
> — Verbatim of the answer, which was provided in English by the student

8.5 Discussion

The aim of this exploratory study was to assess several facets of the interaction with an EAT over a longer period of time, namely two thirds of a semester. In this regard, two classes of a blended university Master about education technology were provided with the possibility to spontaneously use the EAT at any time they wished during periods of distance learning in an entry-level course about programming and computational thinking. The tool was meant to foster learners' self-reflection about their own emotional experience in a course that is known to put students on an emotional roller coaster: students often alternate success in having their intentions translated into code with failures in reaching a final product that meet their overstated expectations (Fritz & Schneider, 2019). In the meantime, the EAT was also allegedly adopted to allow students to project themselves socially and affectively in a course whose instructional design does not impose collaboration between peers. The asynchronous and non-collaborative settings of the course, combined with the overall dexterity of the students with web applications, was therefore considered an interesting scenario to assess the longitudinal use of an EAT with the characteristics advocated in the thesis.

The overall use of the EAT has nevertheless been scant, which has caused a chain effect on all the other indicators linked to the research questions set forth. The discussion will therefore start from this consideration, assessing the first research question, and progressively unravels the other questions successively.

8.5.1 Scarce Use of the EAT

The number of expressed emotions, used here as the sole indicator of concrete use of the EAT, showed that only one student expressed over 80 emotions, with a dozen of other students expressing between 34 and 10 emotions, and the reminder 10 emotions or less, with 3 students not expressing any emotion at all. Even without an accurate benchmark at disposal, these measures refer of a scarce overall adoption of the EAT. The possibility of having students only monitoring the emotional experience of others, without entering any emotions themselves, seems unlikely and did not transpire from any comment.

Standing on indicators of *pure* use of the EAT, it may therefore be assumed that the effort to recall its existence, open, produce and peruse emotional information outweighs the benefits that could be derived from a stable and persistent adoption of the tool (Dillenbourg et al., 2016). The course as a whole, though, is very demanding in terms of different environments, tools, requirements for the projects to be handed-in, documentation to be consulted, and so on (Fritz & Schneider, 2019). As a consequence, there may not even have been a deliberate assessment of the trade-off between costs and benefits, since the overwhelming tension of articulating all the different parts of the course could have simply taken the upper hand.

Another possibility is that, despite some learners having wished to dispose of the emotion of others as in Gaëlle Molinari et al. (2016), others can be reluctant to the very idea of emotional awareness altogether, as encountered by Ruiz et al. (2016). A closer look at the perception of usefulness can therefore provide a more accurate representation of students attitude towards the EAT.

8.5.2 Perceived Usefulness Is Promising at First, but Drops with Concrete Use

The upfront expectations of students towards the EAT seem to relate a genuine need for a richer experience during remote periods, which is consistent with theoretical and practical efforts to improve socio-affective support in distance learning (Brackett et al., 2019; Järvenoja et al., 2013; Kreijns et al., 2013; Lavoué et al., 2020). More specifically, learners manifested the highest expectations about the usefulness of an EAT for the *Understanding Others* ($M_{total} = 6.93$), and *Self-Other Comparison* ($M_{total} = 7.70$), that is, the two more socially-oriented dimensions. In particular, it is worth pointing out that the *Self-Other Comparison* dimension did not even receive any low score in the *expectancy* survey, and the *Understanding Others* only two very low scores (see 8.5). This seems to rule out the possibility that there are students *a priori* reluctant to the whole business of emotional awareness.

Conversely, for the two more self-oriented dimensions, *Self-Understanding* and *Self-Regulation*, expectations were only moderate to low ($M_{total} = 4.57$ and $M_{total} = 4.23$ respectively). In other words, the intra-personal function of an awareness tool is perceived as potentially less intriguing compared to the inter-personal stance (Eligio et al., 2012; A. H. Fischer & Manstead, 2016; Rimé, 2009; Van Kleef, 2018). It would be interesting to investigate whether this is a shortcoming that is overtly attributed to the concept of an EAT, or it also (or rather) subsumes an intrinsic difficulty of students to engage in emotion self-reflection and self-regulation (Lavoué et al., 2020; Gaëlle Molinari et al., 2016; Seager, 2002).

The *Affordance* dimension was rated high ($M_{total} = 6.57$), suggesting that the tool is expected to serve its function of prompting learners to share their emotions. This seems to corroborate the need for a dedicated tool targeting specifically emotions (Brackett, 2019; Brackett et al., 2019; Brackett et al., 2012a), but does not implicate the benefit of implementing an emotional structure into it, since ratings with the *Demo* survey did not increase the expectations.

Finally, the *Frequency* and *Social Presence* dimensions are more difficult to assess, since they are the two dimensions on which the two classes have a lesser convergence. Frequency of use is notoriously a user-experience dimension that is often overestimated by potential users, for whom telling *yes, I will use it* has no concrete consequences (Sauro & Lewis, 2016; Tullis & Albert, 2013a). The moderate expectations for *Social Presence* ($M_{total} = 5.83$), compared with higher expectations for the more socially-oriented dimensions, may suggest that learners consider the EAT as a potential source of social reference rather than a way to perceive the presence of others. This is consistent with theoretical evidence suggesting that emotions in others can be a valuable source of information, especially in challenging situations (Järvenoja et al., 2013; Van Kleef, 2009, 2018).

Inter-dimensional comparison should nevertheless take into account the tentative nature of the EAU scale. Direct comparison between different dimensions, in fact, presupposes that all the seven constructs can be mapped into a similar space (Williams, 2021), which is an assumption that cannot be sustained without a thorough validation of the scale. Intra-individual change over time on the same dimension, though, is less concerned by the *absolute* measurement space of a construct, for it is determined by the relative shift in perception as the situation evolves (Fitzmaurice et al., 2011).

In this regard, all the expectations were basically confirmed in the *Demo* survey, once participants had tested the concrete use of the EAT, suggesting that the specific features of the tool neither increased, nor decreased the expectations about its usefulness. This is both good and bad news. On the bright side, learners seem to trust the tool's ability to sustain their expectations. On the other hand, though, the tool failed to change participants' mind on the dimensions rated moderate to low. This seems to indicate that they did not see in the tool additional interest, especially for interpreting, reflecting on, and experiencing their own emotional episodes (Boehner et al., 2007; Lavoué et al., 2020; Ruiz et al., 2016).

The perception of usefulness then drops immediately. Already at the *halfway* survey, the rating on all dimensions of the EAU have plummeted. Almost none of them even slightly recover at the *final* rating. All ratings tended to stabilize on the bottom end of the scale, so that the dimensions that suffered from the greatest decrease are the ones that benefited from greater expectations upfront, namely *Understanding Others* and *Self-Other Comparison*, who dropped of more than 2.5 rating points.

A similar drop was observed for the *Social Presence* dimension, which started though from only a moderate expectation. This means that the tool was eventually perceived as not conveying the presence of others at all. This may be due, in large part, to the fact that students almost never crossed each others online at the same time. An attempt to reconstruct asynchronously the presence of others through non-contextualized emotional traces left in the tool seems therefore rather hopeless. Some students overtly reported in the open-ended questions that they disposed of other tools to interact with colleagues, such as instant messaging, which provided a more holistic form of interaction, providing cognitive, social and affective support at the same time (Blunden & Brodsky, 2020; Cheshin et al., 2011; Parkinson, 2008). This directly questions the interest of disposing of a dedicated tool, even when it implements an emotion structure that would be difficult to obtain in general purpose environment as an instant messaging application.

Finally, the assessment of the perception cannot ignore that many students reported that they wished they could have used the tool more often, but did not manage to do that because of the overall pressure in meeting deadlines with a functioning project, which is a primary concern for novices in programming (Fritz & Schneider, 2019; D. M. C. Lee, Rodrigo, Baker, Sugay, & Coronel, 2011). The *true* perception of the usefulness of an EAT may therefore not be as bleak as the one depicted by the EAU ratings after the concrete (non-)use of the tool. In this regard, the scale in itself showed promising sensibility and reliability indicators, suggesting also the presence of a three-dimensional structure with dimensions related to the user experience (*Frequency* and *Affordance*), to the intra-personal function of emotional awareness (*Self-Regulation* and *Self-Understaning*), and to the inter-personal function (*Self-Other Comparison*, *Understanding Others*, and *Social Presence*).

8.5.3 Inconclusive, But Promising Settings to Introduce Emotional Competence in the Picture

The use of an appraisal-driven EAT is closely linked to the concept of emotional competence from a componential perspective (Scherer, 2007), as illustrated in Section 8.2.3. The introduction of the Geneva Emotion Competence (GECo) test was therefore a promising measure upon which explore the potential role of this construct as an intervening factor in determining the use and perception of the EAT. Given its accuracy in determining four sub-competences (recognition, understanding, regulation and management of emotion), the length of the test discourage most of participants from taking it, especially during an intensive period as the first semester of the Master. As a consequence, only $N = 11$ students completed the test, which limited the information what could be retrieved from this valuable source.

As already mentioned, the test is targeting more specifically working settings, but the item could be reworded relatively easily to match academic settings, both at the intra-personal and inter-personal level. The accuracy of the test may also be relaxed depending on the specific aims of the study, so that it may therefore be possible to envisage a shorter version. For instance, the recognition part based on efferent manifestations could be skipped in computer-mediated learning environments who do not aim at reproducing face-to-face interactions (Buder, 2011; Janssen & Bodemer, 2013).

All things considering, a performance-based test about the construct of emotional competence may be a very useful instrument in investigating emotional awareness more broadly. Under the assumption that emotional competence may also be trained, the longitudinal use of an EAT under more pro-active conditions can even contribute to test potential effect on the overall competence (Brackett et al., 2012a; Hoffmann et al., 2020).

8.5.4 Questionable Procedure for Testing Emotional Experience From the EAT

Under the assumption that emotional stimuli have a privileged access to memory (Brosch et al., 2013; Kensinger & Ford, 2020; Montagrin et al., 2013; Pool et al., 2015; Rimé, 2005), it has been tested whether students had an accurate recalling of the emotional experience as it was depicted by the EAT at the end of the semester. Even though the practice of recalling self-emotional experience and projecting emotional experience of others have been adopted in other contributions (Avry, Chanel, et al., 2020; Eligio et al., 2012; Gaëlle Molinari, Chanel, et al., 2013), in the present study the assessment was prompted more from enthusiasm than careful planning. Even in the case of a more consequential use of the

EAT, though, doubts may persist about the validity of the procedure. As a consequence, this research question will not be discussed any further, being only slightly related to the overall aims of the thesis.

8.5.5 Consistency Between Classes Suggests Limited Interacting Effects

Finally, a last research question spanned across the previous four and concerned difference between two classes observed during two different academic years, but under very similar settings. It was assumed that in such conditions, at least from an exploratory perspective, the two classes could be retained as a form of pseudo-random assignment, since the composition of classes in the Master MALTT are not usually subject to any systematic factor at play in determining age, background, or perspectives of students (Pearl & Mackenzie, 2018a, 2018b; Rohrer, 2018). It was posited that, if *wide* differences in the use and perception of the EAT between classes could be observed, that would probably be accountable, in part at least, to interacting effects between students of each cohort (Dillenbourg et al., 2016).

Most of the observed measures in this study, though, are very consistent across the two classes. In both classes, only a handful of students adopted the tool over the longitudinal period, with similar patterns in periods where emotions were expressed and periods were students were most likely working on other courses. In Class 1, though, there was a single student that expressed most of the emotions, whereas in Class 2 the encoding of emotional information was less skewed, even though far from being homogeneous. Notwithstanding this difference, the rating on the EAU scale over the four surveys was almost symmetrical between the two classes. The dimensions that have been rated somehow differently were limited to the *Expectancy* and *Demo* surveys and consisted in *Frequency*, where Class 1 was more optimistic; *Affordance,* once again with Class 1 having a more favorable outlook; and *Social Presence*, where Class 2 reported greater expectations. The fact that these differences concern ratings before the concrete use of the EAT, though, rule out the possibility that they may be determined by interacting factors, at least not implicating the use of the tool. As a reminder, classes spent the first of the three periods during the semester without using the EAT. So, different social interactions may have occurred across classes that could, in principle, have influenced students in having different attitudes toward the use of an EAT with their colleagues. The use of the tool could be implemented from the very onset of the semester to limit this possibility.

Inter-individual attitudes towards the EAT were nevertheless present, with some students having adopted it, or having at least a favorable outlook towards its usefulness, as indicated by the open-ended questions. Those students have nevertheless not *pushed* others in following through, and were most probably even set back from the lack of use from the colleagues. In this sense, there may have been an interacting effect, with potentially interested students being dissuaded by their isolated use of the EAT. As already mentioned, this would in fact be a reverse effect, with the present of the EAT increasing isolation rather than allowing students to project themselves socially and affectively in the computer-mediated environment.

On the bright side, the consistency between the two classes at least conveys support for the intrinsic quality of the EAU scale and also of the EAT in itself. The EAU scale seems to have allowed both classes to consistently express their expectations and progressive disillusion with the usefulness of the tool. Furthermore, notwithstanding the limited use, the rating on the SUS (Brooke, 1996) was consistent between the two classes, yielding an overall results of almost 73 points. This last rating may suggest that there has not been a favorable interacting process between the intrinsic quality of the tool and the conditions for taking advantage from it.

8.6 Conclusion

The present observational and longitudinal study aimed at investigating the use and perception of usefulness of an EAT during asynchronous and non-collaborative learning settings, under the assumption that the presence of the EAT may provide instrumental information on various dimensions that are considered as pivotal in remote computer-mediated learning environments. Furthermore, by implementing the EAT in two cohorts of the same course, the study also aimed at investigating whether differences between the two classes could be observed. If that would have been the case, the use and perception of the EAT may depend also on interacting dynamics between learners, rather than on individual characteristics alone. Finally, as a corollary objective, the study also attempted at determining whether emotional competence may emerge as an intervening factor in the use or perception of the EAT.

Overall, the study presents several drawbacks that limited its informative potential. At the same time, it also provided some promising elements, particularly in the material and type of analysis which may help future investigations about the subject at hand. The concluding remarks of this chapters first assess the main limitations of the study, and then resume the main contributions that may be considered for future studies.

8.6.1 Limitations

The study suffered from an unresolved ambiguity between a field study and a pedagogical intervention into an ongoing course. On the one hand, the field study objectives presupposed minimal intervention, consistently with a *pure* awareness tool perspective, letting learners free to adopt and use the tool in the way they saw fit. At the same time, from a pedagogical standpoint, one of the established shortcoming in the implementation of auxiliary technologies in the learning activity is that their presence alone does not guarantee neither the use, nor their effectiveness (Janssen et al., 2011; Kreijns et al., 2003, 2003). The overall sparse use of the EAT, consistent between classes, has therefore diminished the informative potential of the whole study, which proved to be exceedingly ambitious in its *group-awareness* perspective. In fact, learners were bestowed with the whole process of reminding to use the tool, open it alongside the different learning activities necessary to comply with the course requirements, and extrapolate meaningful information from displaying their own emotions, as well as monitoring their emotions and that of their colleagues. As many learners have revealed in the free commentary, this was an excessive demand in a course where they already had to face new and complex learning environments or tools

(Fritz & Schneider, 2019).

A certain amount of *scripting* – that is, pedagogical scaffolding of the learning activities (Miller & Hadwin, 2015) – seems therefore necessary. Learners should be more adequately supported in the use of the EAT, for instance with occasional reminder of its presence. Discussions about its use and the available information could also been integrated, for example in the form of focus-groups (Lallemand & Gronier, 2017). This sort of support may be very consistent with an interacting approach. The use of the EAT would be assessed and reassessed by the group as a whole, with guidance resulting as a direct effect of the input from students. With hindsight, the neutral comparison between cohorts was therefore overly ambitious, for it requires a *laissez-faire* attitude in contrast with learners needs of a more guided scenario and support. This is even more relevant considering the convenience sample of the study: learners that by vocation and training are exposed to various learning technologies. The fact that even learners allegedly open to adopting new technologies made a sparse use of the EAT suggests that this design is unlikely to provide interesting evidence on the subject at hand, beyond confirming that an awareness tool must be integrated in an environment carefully planned to maximize its adoption and efficiency (Buder, 2011; Janssen et al., 2011; Kreijns et al., 2003, 2003).

Beside the context in which the EAT has been deployed, also the tool in itself presented major limitations from a longitudinal perspective. Even though the overall perception of the tool on the *Demo* survey and the System Usability Scale (Brooke, 1996) do not highlight that students perceived limitations, they were also not in the full condition to assess its potential. The use of three word clouds to peruse emotional information is a very limited representation of the complexity and quality of emotion as an information unit, a central tenet of the overall contribution. The conceptual work done on implementing an emotional structure into the tool was therefore limited to the expressing-displaying part of the EAT, but was not then reflected in the perceiving-monitoring part. This certainly calls for intervention in conceiving and deploying better graphical representations of emotions, which may walk students along a rich and meaningful representation of their own as well as their colleagues emotions (Berset, 2018; Fritz, 2016b; Leony et al., 2013; Ruiz et al., 2016).

8.6.2 Study's Overall Contribution

Despite all the limitations, the study also provides some elements of interest that may be adopted in future studies about emotional awareness in computer-mediated learning environments more generally.

First, the EAT did not encounter any technical problem. Even though it hasn't been put really under pressure from participants, it has nevertheless been up and running for the two longitudinal studies, apparently collecting all the relevant data provided by students. The intrinsic quality of the tool in itself has not been diminished by the lack of its use, corroborating the idea that the sparse use could be, at least partially, due to the design of the study. The perceived usability is in fact consistent across classes and similar to a previous usability test conducted in synchronous and collaborative settings (Fritz, 2015). The EAT's usability will be more thoroughly assessed in the next chapter, but it is worth mentioning here that it appears the tool can be implemented also in an asynchronous and individual use, without modifying the perception of its usability.

Second, the tentative Emotional Awareness Usefulness (EAU) scale showed promising results with respect to its sensitivity and reliability. The scale nevertheless needs further development, since even the formulation of items considered the vantage point of participant's knowledge. Furthermore, there are at least a few dimensions that are missing and some of the current ones could be decoupled to avoid some ambiguity. For instance, the *Social Presence* item does not take into account the full complexity of the construct as illustrated by the extant literature on the subject (Jézégou, 2010; Kirschner et al., 2015; Lowenthal & Snelson, 2017; Rourke & Kanuka, 2009), see also Section 1.3.3. Social isolation is in fact only a facet of social presence. For example, the scale does not include anything about a sense of belonging or authenticity, which are characteristics advocated by several scholars. The *Affordance* dimension suffers from the very same ambiguity that the thesis as a whole tries to dissipate, since it only consider the *self-affordance* in expressing one's own emotions, but does not consider the *social-affordance* in considering others' emotions. The *Understanding Others* and *Self-Other Comparison*, in fact, go a step further and implicitly assume that others' emotions are taken into consideration, whereas this dimension should be indexed in its own term. One can in fact be interested in others' emotions without necessarily acting on their understanding, or do not use the emotions' in others as a comparison, but for other purposes. Furthermore, the scale could also be deplyoed in collaborative settings, in which case *Self-Regulation* should also be supplemented by an item on inter-personal emotion regulation (Reeck et al., 2016; Zaki & Williams, 2013). In other words, the scale itself should be better tailored on the abstract model of emotional awareness presented in Section 2.5. Finally, the scale should seek a better integration with existing scales in the field (Kreijns et al., 2011; Lane et al., 1990; Pekrun et al., 2011, 2016; Rieffe et al., 2008; Tu & McIsaac, 2002), especially in the perspective of having multiple items about the same dimension. Considering the growing interest towards the instrumentality of emotional awareness in computer-mediated learning environments, it would be worth investing in the development of a scale that maintains the general outlook, but that better reflects good practices and standards in the development of quality measures (Boateng et al., 2018; Flake & Fried, 2020).

Finally, speaking of valid instrument, the Geneva Emotional Competence (GECo) test (Schlegel & Mortillaro, 2018) provides a viable option for researchers interested in in an ability measure of this socio-affective construct. Even if the test is quite long, it provides measurement for emotion recognition, emotion understanding, emotion regulation, and emotion management, which can be related each to different facets of the use and perception of an EAT. Furthermore, the items used outside emotion recognition are situated in working conditions that have a bearing with learning in higher education. It therefore presents also some ecological potential, maintaining though the rigor of a scientific instrument.

All things considering, thus, also this second empirical contribution is consistent with the thesis's general objective of focusing primarily on construct validity and internal consistency. The generalization potential and practical evidence yielded by the study are very limited, but the promising quality of the individual instruments adopted may be of interest for future studies wishing to take into account the longitudinal effect of emotional awareness.

8.6.3 Acknowledgments

I would like to thank professor Daniel K. Schneider for allowing the use of the EAT in two occasions during the STIC I course.

Chapter 9

A Data-Driven Assessment of the Emotion Awareness Tool in Different Computer-Mediated Environments

This chapter provides a general assessment of some of the fundamental features of the Emotion Awareness Tool (EAT) implemented in the thesis from a data-driven perspective. To do so, the study will merge data from the usability test (Fritz, 2015), briefly resumed in Section 4.3.4, and the two empirical contributions of Chapter 7 and Chapter 8. The structure of the chapter will nevertheless follow once again the traditional structure of an empirical contribution (Sollaci & Pereira, 2004) already adopted by the two previous empirical contributions.

9.1 Study Rationale

Even though the use of an EAT is highly dependent on a series of factors such as who is using it, when, and why, one of the main tenet of the thesis is that it is possible to provide a multi-purpose tool, which can be adapted to different situations. Taking advantage from the fact that the EAT has been deployed in three empirical settings presented in this contribution – the usability test retrived from Fritz (2015) and the two previous empirical chapters – data collected in each occasion can be integrated to provide a more thorough data analysis.

More specifically, it is possible to distinguish between the synchronous and collaborative (Synch./Collab.) settings of the usability test and the experiment in Chapter 7 on the one hand, and the asynchronous and individual (Asynch./Indiv.) – or non-collaborative – settings of the study in Chapter 8 on the other. By this comparison, though, the aim is certainly not to draw inferences from the effect of different computer-mediated environments on the use of an EAT (*e.g.,* students are *happier* working together than in isolation). On the contrary, consistently with the focus on internal validity of the thesis, the objective is rather to assess whether the same EAT can adapt to different settings. One of the central tenet of the thesis is, in fact, that researchers and practitioners interested in endowing computer-mediated learning environments with emotional awareness may do so by adapting a tool, whose central features may be of interest in a wide number of scenarios. At the same time, the EAT is also based upon specific theoretical and technical assumptions, such as the usefulness of implementing an emotion structure into the tool. A comprehensive analysis of data gathered in different settings, but based upon similar theoretical and technical postulates, can therefore represents a first assessment of the intrinsic qualities and shortcomings of the EAT.

9.2 Research Questions

In this regard, the chapter proposes four types of assessment related either to specific features of the EAT or its perception as a whole. Each assessment focuses on a specific research question, presented hereafter.

9.2.1 Appraisal Dimensions as Meaningful Evaluation of Events

The concept of appraisal is a central tenet of the entire thesis. It is the *glue* that puts together different theoretical assumptions, such as what an emotion *is* (Moors et al., 2013; Sander, 2013; Scherer, 2005, 2019a, 2022), and pedagogical implications, such as how learners can take full advantage from an EAT in terms of self-reflection and strategic communication of their emotional experience to others (Boehner et al., 2007; Cernea & Kerren, 2015; Lavoué et al., 2020; Gaëlle Molinari, Chanel, et al., 2013). Appraisal is technically translated into the interface with the use of sliders, upon which learners can evaluate the situation in a dimensional approach to emotion self-report. In the three empirical settings in which the EAT has been adopted, the *Valence* and *Control/Power* dimensions represented the appraisal criteria, which prompted a cognitive evaluation of the situation by learners. The first research question therefore relates to the use of the sliders as representative of the interest and instrumentality of the underlying appraisal dimensions.

9.2.2 Lexicalized Emotions as Representative of Learners' Subjective Feelings

Another focal point concerns the symbolic representation of the whole emotional experience that learners can use to extrapolate inter-personal meaning making and/or send a strategic *token* that maximize inference of the *true* emotional episode in others (Fontaine et al., 2013; Grandjean et al., 2008; Ogarkova, 2013). In the EAT, this role is conferred to a number of lexicalized emotions pre-compiled in the underlying affective space, which are meant to provide learners with meaningful options to coalesce the emotional experience in a practical and intuitive way. The second research questions therefore assess to what extent the lexicalized emotions proposed in the EATMINT circumplex met this requirement.

9.2.3 The Computational Model on Trial

The core of the EAT relies on the computational model presented in Chapter 5, whose fundamental tenet is that the holistic emotional experience can be predicted on the basis of the cognitive evaluation of the situation (Fontaine et al., 2021; Gentsch et al., 2017; Scherer & Fontaine, 2018; Scherer & Meuleman, 2013; Scherer et al., 2018). The EAT harness this causal mechanism to provide a subset of lexicalized emotion that are most likely to represent learners subjective feeling, given the rating on the appraisal dimensions. In all three empirical settings, the EAT provided 3 suggestions as buttons, but also left participants free to provide another emotion term either from the pre-compiled list (but not proposed as button) or outside the list. The third research questions investigates the efficacy of the computational model in providing learners with an *educated guess*, which was accepted as an accurate representation of learners' subjective feeling.

9.2.4 Perceived Usability Beyond the Concrete Use

Even though the concrete use of an EAT is the primary concern for an awareness tool, it may be influenced by a number of dispositional or situational factors, as illustrated by the abstract model of emotional awareness in Section 2.5. In other words, the EAT may have a *potential* that is not fully expressed in the situation at hand. The perceived usability of the tool – which comprises the efficacy, efficiency and satisfaction in using the EAT (Brooke, 1996; Lewis & Sauro, 2018; Tullis & Albert, 2013a) – may therefore complement performance-based data in the assessment. The fourth research questions thus refers to how the usability of the EAT has been rated across different settings.

9.3 Methods

This third empirical contribution can be considered a form of secondary data analysis (Weston et al., 2019) or a small internal meta-analysis of one's own studies (Goh et al., 2016). It consists in grouping the datasets of the three empirical contributions where the EAT has been deployed: the usability test in Fritz (2015), the experiment of Chapter 7, and the longitudinal study of Chapter 8. Data will be integrated in two datasets, which will be used to address the four research questions.

9.3.1 Expressed Emotions Dataset

The first dataset consists in all the emotions that have been expressed through the EAT. This dataset comprises $N_{observations} = 1097$ produced by $N_{participants} = 75$. Table 9.1 illustrates how observations and participants are divided between the three datasets, with two datasets attributed to the Synch./Collab. setting, and one to the Asynch./Indiv. setting.

Table 9.1: Descriptive statistics in the rating of the two appraisal dimensions of the affective space

Dataset	Setting	Participants	Observations
Usability Test	Synch./Collab.	14	240
Chapter 7	Synch./Collab.	35	483
Chapter 8	Asynch./Indiv.	26	374

9.3.2 System Usability Scale Dataset

The second composed dataset comprises the rating on the System Usability Scale (Brooke, 1996), which was administered in the usability test and in the study of Chapter 8, but not in the experiment of Chapter 7. In total, the SUS has been rated by $N = 40$ participants, 14 in the Synch./Collab. setting of the usability test, and 26 in the study of Chapter 8.

9.3.3 Analyses

The analyses will be comparative in nature, but without the use of inferential statistics, which are considered outside the scope of a preliminary assessment of the tool.

9.4 Results

The results are presented with respect to the four research questions. A thorough use of graphical representations is adopted to better convey the different facets of the measures at hand (Franconeri, Padilla, Shah, Zacks, & Hullman, 2021; Hegarty, 2011; Tukey, 1977).

9.4.1 Measures About the Use of Appraisal Dimensions

The first theory-driven feature of the EAT under scrutiny are the two sliders, which represent the appraisal dimensions through which the eliciting event is evaluated (Scherer, 2001, 2005, 2009b). As a reminder, the EATMINT circumplex adopts the *Valence* and *Control/Power* appraisal dimensions to prompt the evaluation of the situation. In all empirical contributions, the Valence dimension was prompted with the question *Is the situation pleasant?* The *Control/Power* dimension was prompted with the question *Is the situation under your control?* in the Synch./Collab. contributions. For the Asynch./Individ. contribution, on the other hand, the question *Can you modify the situation?* was adopted instead. Both dimensions could be rated from a negative pole labeled *Not at all*, corresponding to a score of -100, to the positive pole labeled *Yes, absolutely*, corresponding to a score of 100. Each slider was sensitive to 1-point variation.

Overall Ratings of the Appraisal Dimensions

One of the interesting indications that can be assessed through the ratings on the appraisal dimensions is to what extent participants could make use of the full range of the slider, that is, whether they discriminate the eliciting events as being more or less pleasant, and more or less under their control. In this regard, Table 9.2 reports the number of participants expressing at least one emotion through the tool, the cumulative number of emotions expressed, as well as the overall mean and standard deviation of the two appraisal dimensions.

Table 9.2: Descriptive statistics in the rating of the two appraisal dimensions of the affective space

Setting	Participants	Observations	Valence	Control/Power
Asynch./Indiv.	26	374	0.06 (59.00)	5.21 (50.07)
Synch./Collab.	49	723	9.84 (56.68)	-3.15 (60.12)
Total	**75**	**1097**	**6.50 (57.64)**	**-0.30 (57.01)**

Results show that, for the Valence dimension, the overall mean is almost perfectly neutral for the Asynch./Indiv. setting, whereas it is slightly positive (around 6 points) for the Synch./Collab. setting. In both settings, the rating of the Valence dimension yielded a high standard deviation of around 60 rating points. Data therefore corroborate that participants in both settings took advantage of the full range of the Valence dimension in a very similar way. With respect to the *Control/Power* dimension, the overall mean for the Asynch./Indiv. setting is slightly positive (around 5 points), whereas it is slightly negative for the Synch./Collab. setting (around -4 points). The standard deviations are also high, but more diverging, with a difference of more than 10 rating points (around 50 for Asynch./Indiv. against around 64 for Synch./Collab.). For this second appraisal dimension, thus, data highlight a slight divergences in central tendency, even though it remains close to the neutral point for both settings, and less variance in the rating of the *Control/Power* slider in the Asynch./Indiv. setting.

The descriptive measures are complemented by Figure 9.1, which shows the density of the two appraisal dimensions for each learning setting. The plots highlight fairly symmetrocal and flat distributions (*i.e.*, Leptokurtic-like shapes) around the neutral point for all combinations, except for the *Control/Power* dimension in the Asnyhc./Indiv. setting, on the top-right plot, which has an higher peek of the distribution (*i.e.*, Platykurtic-like). This higher peek is nevertheless inflated by a single participant who expressed 65 emotions leaving the *Control/Power* dimension on the neutral point. The distributions also denote some *bumps* on the tails, especially in the positive tail of the Asynch./Indiv. and, to a lesser extent, both tails in the Synch./Collab. setting for the *Control/Power* dimension. These *bumps* represent ratings in which participants used the extreme poles of the sliders. There is therefore a certain trend in expressing the more extreme values on the appraisal dimensions.

All things considering, though, participants in both settings seem to take advantage, individually, of the full range of the appraisal dimensions. Being the two dimensions related, though, the analysis must also consider their joint ratings, which is presented next.

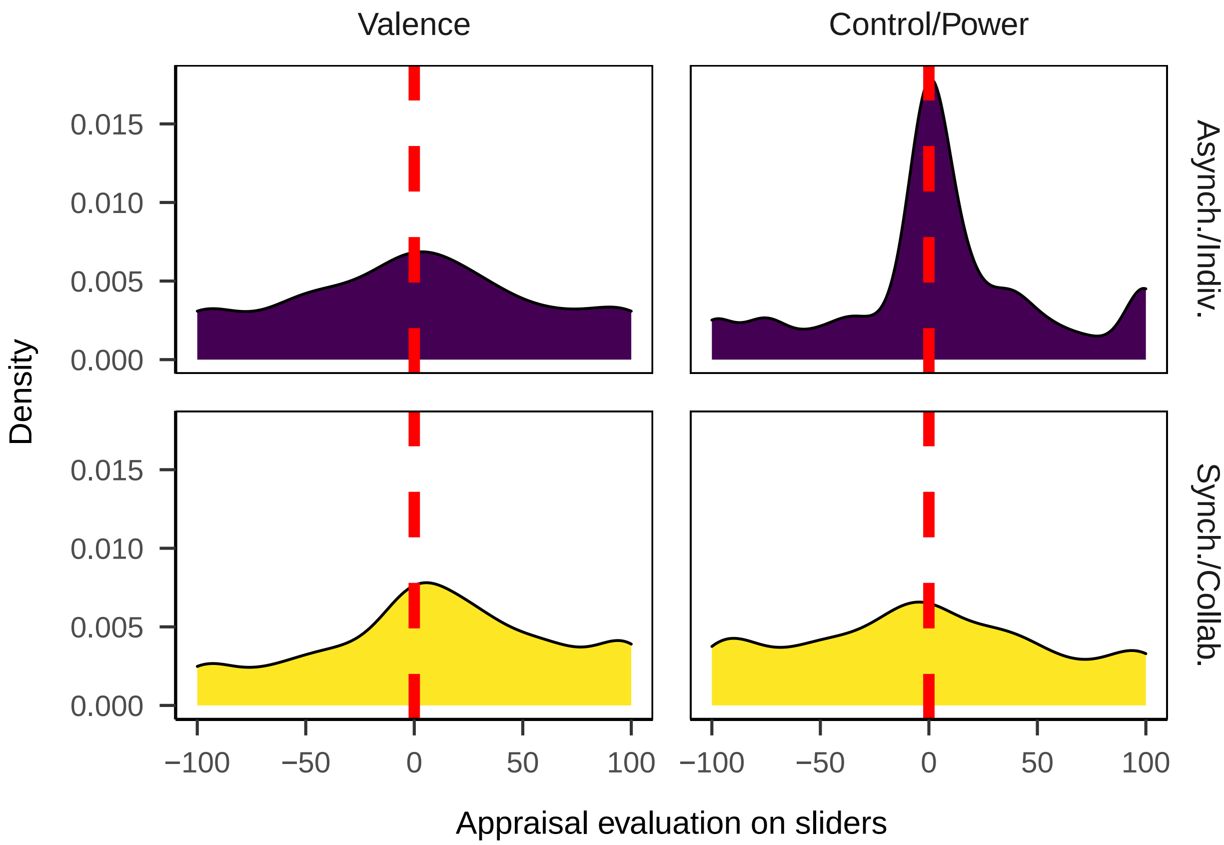


Figure 9.1: Density plots of the two appraisal dimensions' ratings for the each setting.

Joint Ratings of Valence and Control/Power

The second element of interest in the use of the appraisal dimensions is whether their use is independent from one another, in which case the dimensions are truly orthogonal, or if there is a sort of *multicollinearity* due to a high correlation between ratings (see Section 4.2.3 for a more theoretical discussion on the subject). In other words, does it happen that participants rate a situation as pleasant, but without feeling control over it; or, conversely, a situation as unpleasant, but feeling control over it? Figure 9.2 shows two Locally Estimated Scatterplot Smoothing (LOESS) functions (Jacoby, 2000) – that is, two non-parametric regressions lines that best fit the data at hand – applied to the points defined by the *Valence* appraisal on the x-axis, and the *Control/Power* appraisal on the y-axis.

The fitted lines highlight a strong positive correlation between the *Valence* and the *Control/Power* ratings, especially in the Synch./Collab. setting, for which the relationship is almost perfectly rectilinear. The ratings of the two appraisal dimensions tend thus to co-vary, so that *Valence* and *Control/Power* are both negative or both positive. This phenomenon is corroborated if the expressed emotions are divided in three possible combinations: (1) appraisal dimensions share the same sign; (2) appraisal dimensions are of opposite sign; and (3) either or both appraisal dimensions are on the neutral point 0. Table

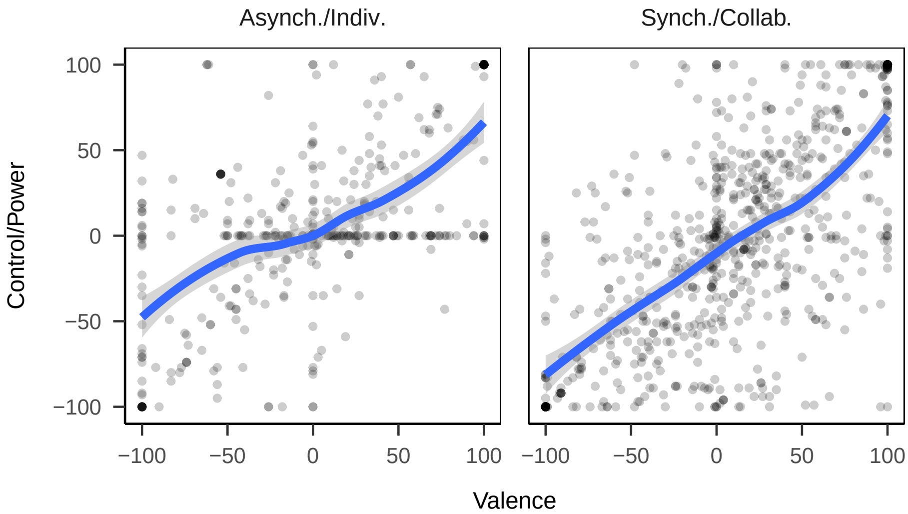


Figure 9.2: LOESS functions applied to *Valence* x *Control/Power* appraisals in the two settings.

9.3 reports the number of participants that expressed at least one emotion with the appraisal combination, as well as the cumulative number and relative proportion of observations.

Table 9.3: Emotions expressed with different combinations of appraisal dimensions.

Combination	Participants	Observations	Proportion
Asynch./Indiv. (26 participants)			
Same sign	25	185	0.49
Opposite sign	19	72	0.19
Either or both neutral	17	117	0.31
Synch./Collab. (35 participants)			
Same sign	48	509	0.70
Opposite sign	43	170	0.24
Either or both neutral	18	44	0.06

For both settings, the same sign combination was expressed at least one time by a greater number of participants, and proportionally more than the other combinations: almost half of the total (0.49) for the Asynch./Indiv. and almost three quarter of the total (0.72) for the Synch./Collab. setting. The opposite sign combination, on the other hand, accounts only for around one fifth of the total both for the Asynch./Indiv. (0.19) and the Synch./Collab. (0.21) settings. Finally, the neutral point was used in greater proportion in the Asynch./Indiv. setting, with almost one third (0.31) of the total, compared to the

Synch./Collab. setting with a proportion of around one in twenty (0.06). As in the case of the density plots described above, this score is inflated by a single participant in the Asynch./Indiv. setting who expressed more than 60 emotions leaving the neutral point on the *Control/Power* dimension.

Individual attitude towards rating both dimensions together should therefore be considered in the assessment. In this regard, it is interesting to compute the individual correlation between the rating of the *Valence* dimension and the *Control/Power* dimension, and then use the average of the correlation (M_ρ) as an indicator of whether the two sliders tend to co-vary or not. An $M_\rho \to 0$ would suggest the two dimensions are orthogonal, whereas $M_\rho \to \pm 1$ would suggest that the two dimensions are rated exactly in the same way (positive correlation), or that one is rated asymmetrically with respect to the other (negative correlation). Among the $N = 74$ participants that expressed at least 2 emotions (the lower bound to compute an intra-individual correlation), the average correlation observed is of $M_\rho = 0.47$ ($SD_\rho = 0.43$). The correlation is greater in the Asynch./Indiv. setting, with $M_\rho = 0.58$ ($SD_\rho = 0.47$) compared to the Synch./Collab. setting, where $M\rho = 0.42$ ($SD_\rho = 0.41$). In both settings, thus, there is a highly positive correlation between the two sliders, which tend to be rated symmetrically.

Synthesis

All things considering, data suggest that participants take advantage of the full range of each appraisal dimensions individually, but, combined, the two appraisal dimensions are not used as orthogonal. On the contrary, there is strong correlation between the two ratings. This phenomena is consistent in both settings. To assess whether this is problematic, though, it is necessary to check for the subjective feelings that have been expressed with the appraisal ratings. In fact, it could be the case that participants predominantly expressed subjective feelings that are theoretically characterized by either positive *Valence* and positive *Control/Power*, or negative *Valence* and negative *Control/Power*, for that would explain the lack of orthogonality. The link between appraisal dimensions and subjective feelings is illustrated below in the chapter.

9.4.2 Measures About the Subjective Feelings Expressed

The second assessment concerns the expression of the subjective feeling, that is, the conscious experience of the emotional episode that is usually labeled using natural language words or idioms. (Fontaine et al., 2013; Grandjean et al., 2008; Ogarkova, 2013). The assessment primarily aims at determining to what extent the lexicalized emotions included in the underlying affective space met participants' need in terms of representation and differentiation of the conscious experience of the emotion (Barrett et al., 2001; Erbas et al., 2015). As a reminder, the EATMINT circumplex proposes 20 lexicalized emotions: Amused, Annoyed, Anxious, Attentive, Bored, Confident, Confused, Delighted, Disappointed, Disgusted, Empathetic, Envious, Frustrated, Interested, Irritated, Relaxed, Relieved, Satisfied, Stressed, and Surprised. If these 20 lexicalized emotions meet learners' need in best describing their subjective feeling, participants should have made a spare use of the possibility to express their feelings with natural language words or idioms falling outside this list. In this regard, it is worth comparing whether the lexicalized emotions of the underlying affective space are consistent with participants needs in the two settings. Table 9.4 illustrates the cumulative number of expressed subjective feelings listed or not listed in the underlying affective space.

Data show that in both settings, participants privileged the lexicalized emotions included in the EATMINT circumplex with proportions above .80. There is nevertheless a difference between the two conditions of more than ten percentage points, since the proportion of listed options for the Asynch./Indiv. setting is around 0.83 against 0.96 in the Synch./Collab. setting. A difference between the two settings is also reinforced by the number of distinct subjective feelings expressed outside the proposed list. In the Asnych./Non-Coll. setting, participants provided 30 distinct subjective feelings, mostly using single words, compared to 12 distinct subjective feelings in the Synch./Collab. setting. These results suggest that, in the Asynch./Indiv. setting, participants may need a richer *emotional vocabulary* (Barrett, 2017; Erbas et al., 2015; Ogarkova, 2013) to express their conscious emotional experience, even though the 20 lexicalized emotions proposed by the underlying affective space cover their needs most of the time.

Another aspect worth considering in the assessment of the subjective feelings is whether the relative frequency in expressing each option listed in the circumplex varies across settings. Table 9.5 reports the relative frequency of each one of the 20 lexicalized emotions in the EATMINT circumplex for both settings, as well as the absolute difference across settings. The use of the absolute difference highlights the fact that this comparison does not aim at determining whether participants in one setting tend to experience a specific feeling more or less than in the other setting, since the two empirical contributions proposed in the thesis are not fit for this purpose. The aim of the comparison is rather to determine whether the same underlying affective space may adapt to different *needs* in conveying the holistic emotional experience.

Data illustrate roughly three different combinations. First, there are options with a high relative frequency in one setting and a low relative frequency in the other (*e.g.*, *Amused*, *Relieved*, *Surprised* or *Satisfied*). Second, there are options with high relative frequencies which are consistent across settings (*e.g.*, *Attentive*, *Interested*, *Bored* or *Frustrated*). Finally, there are feelings with low relative frequencies in both settings (*e.g.*, *Envious*, *Disgusted*, *Relaxed*, *Irritated*, *Empathetic*). Results therefore corroborate the assumption that the EATMINT circumplex may adapt to different settings, even though some of its lexicalized emotions are seldom chosen by participants. Whether these feelings should continue to be proposed as choices in the affective space is discussed below.

Table 9.4: Cumulative number of subjective feelings expressed that were listed or not listed in the underlying affective space

Setting	Observations	Not Listed	Listed	% Listed
Asynch./Indiv.	374	62	312	0.83
Synch./Collab.	723	29	694	0.96
Total	**1097**	**91**	**1006**	**0.92**

Table 9.5: Relative frequencies of the 20 listed subjective feelings and absolute differences between settings

Feeling	Total	Asynch./Indiv.	Synch./Collab.	|Difference|
Quadrant I. Positive Valence x Positive Control/Power				
Confident	4.97	5.77	4.61	1.16
Interested	10.44	11.54	9.94	1.60
Amused	12.62	2.88	17.00	14.12
Delighted	4.17	5.13	3.75	1.38
Attentive	10.83	11.22	10.66	0.56
Quadrant II. Positive Valence x Negative Control/Power				
Satisfied	3.38	7.69	1.44	6.25
Relaxed	1.79	1.60	1.87	0.27
Surprised	4.87	0.32	6.92	6.60
Relieved	4.27	10.58	1.44	9.14
Empathetic	1.39	0.00	2.02	2.02
Quadrant III. Negative Valence x Negative Control/Power				
Confused	10.04	5.13	12.25	7.12
Anxious	3.38	4.81	2.74	2.07
Bored	7.06	8.33	6.48	1.85
Stressed	4.08	6.73	2.88	3.85
Disappointed	5.47	6.41	5.04	1.37
Quadrant IV. Negative Valence x Positive Control/Power				
Frustrated	6.56	8.33	5.76	2.57
Envious	0.30	0.00	0.43	0.43
Disgusted	0.89	0.64	1.01	0.37
Annoyed	1.49	1.92	1.30	0.63
Irritated	1.99	0.96	2.45	1.49

9.4.3 Measures About the Computational Model Linking Appraisal Dimensions and Subjective Feelings

The two previous sections highlight that, separately, the appraisal dimensions and the lexicalized emotions composing the affective space seem to adapt to the two different settings. The core of the DEW is nevertheless the theory-driven, computational link between the appraisal dimensions and the subjective feeling. It is therefore pivotal to assess whether the parsimonious computational model that suggests a subset of subjective feelings given a specific evaluation of the eliciting event eases learners' task in conveying the holistic emotional experience. The link between appraisal dimensions and subjective feeling can be derived only for the emotions that are part of the underlying affective space, and therefore the following analysis will filter out subjective feelings not included in the EATMINT circumplex list.

The link between appraisal dimensions and subjective feelings can be assessed mainly in two ways. The first is pragmatic, and pertains to the actual use of the tool with respect to the frequency by which learner's chose one of the options proposed in the subset of buttons on the interface, rather than having to recur to the drop-down menu or typing the response themselves. The second is more theoretically-driven and consists in comparing the underlying affective space – that is, the one *expected* from the theory – with the *observed* affective space, which can be computed using the means of *Valence* and *Control/Power* every time a given lexicalized emotion has been chosen by learners.

Frequency of Choice of the Proposed Subjective Feelings

The pragmatic assessment consists in computing the frequency by which learner's *accepted* to click on one of the three proposed buttons labeled with a lexicalized emotion, given the evaluation provided through the two sliders representing the appraisal dimensions. In other words, the frequency represent the number of times that learners found one of the suggested options as the *right* representation, or *best* approximation, of their subjective feeling. The frequency can therefore range from 0 – that is, the learner never found the corresponding subjective feeling in the buttons and had to provide it through the drop-down menu or text input – to 1, in which case the learner always *accepted* one of the three suggestions provided by the buttons.

This kind of measure, though, can be influenced, among other things, by (1) the sheer number of emotions expressed, with low numbers inflating either the opposite poles or the central tendency; and (2) individual characteristics such as conformity to accept a suggestion or the dexterity in choosing another feeling from the list. For these reasons, the frequency is computed first computed individually for each participant that has expressed at least five emotions, and then averaged over all participants in the same setting – that is, $N = 14$ in the Asynch./Indiv. setting, and $N = 47$ in the Synch./Collab. setting. Results are shown in Table 9.6.

Table 9.6: Frequency of clicks on one of the suggested subjective feelings

condition	n	mean	sd
Asynch./Indiv.	14	0.70	0.36
Synch./Collab.	47	0.79	0.32
Total	**61**	**0.77**	**0.33**

Overall, the frequency of clicks on one of the suggested options is $M = 0.77$ ($SD = 0.33$), which means that around 4 out of 5 emotions are expressed using one of the lexicalized emotions suggested through the three buttons on the interface. In the Asnyhc./Indiv. setting, the average frequency is $M = 0.70$ with a standard deviation of $SD = 0.36$. In the Synch./Collab. setting, the average frequency is $M = 0.79$, with a standard deviation of $SD = 0.32$. There is nevertheless a difference of around 0.2 points between the two settings, with participants in the Synch./Collab. setting clicking more often on the buttons. The data, pictured in Figure 9.3, reveals that most of the participants in the Synch./Collab. setting always *accepted* one of the proposed subjective feelings, whereas in the Asynch./Indiv. setting there is a more heterogeneous disposition. This is consistent with evidence in the previous sections of the chapter indicating the need of a more nuanced emotional expression in the Asynch./Indiv. setting.

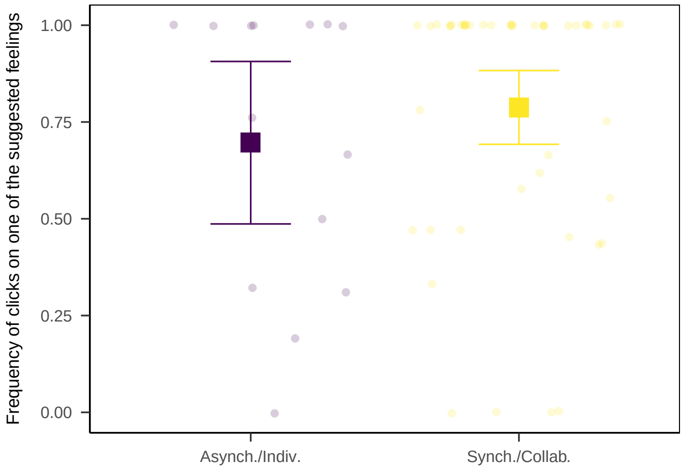


Figure 9.3: Frequency of click on one of the three buttons labeled with a subjective feeling. Bars represent 95% confidence intervals.

Overall, though, the parsimonious computational model fitted into the EAT seems to adequately connect the appraisals dimensions with the subjective feeling: participants took advantage of this feature on average in four out of five emotions expressed. These results seem also to corroborate the limited number of suggestions proposed by the tool. Three buttons, in fact, seem to provide learners with sufficient options to discriminate their subjective feelings. The sheer frequency, though, does not guarantee that this mechanism *works* in the same way at every level of combination between the appraisal dimensions, for which a more detailed analysis is necessary.

Expected Versus Observed Affective Space

With the data collected every time a subjective feeling is expressed, it is possible to compute an observed position of the lexicalized emotion on the circumplex. The observed position is computed in two steps. First, the means of the *Valence* and *Control/Power* dimensions are calculated for every feeling belonging to the underlying circumplex. For example, every time that the subjective feeling *Attentive* has been expressed, the corresponding ratings that the participant has made of the two appraisal dimensions are pooled to determine the means. Once the mean of *Valence* and *Control/Power* are obtained, they are injected into the computational model to retrieve the slope, which will be used to retrieve the *average* angle of feeling on the circumplex. Table 9.7 reports the necessary figures to compute the observed slope and compares it with the expected slope, that is the position of the feeling on the EATMINT circumplex. The absolute difference between the two slopes is also provided. The greater the absolute difference, the wider the gap between the *theoretical* position proposed by the underlying affective space and the *empirical* rating made by participants.

Table 9.7: Aggregated means of appraisal ratings for each of the 20 subjective feelings in the EATMINT circumplex, with observed and expected slopes.

Feeling	N	Valence	Contr./Pow.	Obs. Slope	Exp. Slope	|Slope|
Quadrant I. Positive Valence x Positive Control/Power						
Confident	49	17.61 (29.02)	42.90 (35.43)	22.32	9	13.32
Interested	105	47.21 (35.14)	45.82 (40.17)	45.86	27	18.86
Amused	127	47.87 (42.03)	33.43 (52.02)	55.08	45	10.08
Delighted	42	71.36 (32.52)	64.17 (37.95)	48.04	63	14.96
Attentive	109	30.88 (37.32)	16.55 (34.80)	61.81	81	19.19
Quadrant II. Positive Valence x Negative Control/Power						
Satisfied	34	59.74 (38.40)	34.62 (42.99)	59.91	99	39.09
Relaxed	18	55.17 (37.53)	11.83 (39.23)	77.89	117	39.11
Surprised	49	27.41 (38.08)	-25.35 (39.15)	132.76	135	2.24
Relieved	42	48.64 (42.71)	5.69 (54.03)	83.33	153	69.67
Empathetic	14	13.29 (30.75)	-40.57 (35.90)	161.87	171	9.13
Quadrant III. Negative Valence x Negative Control/Power						
Confused	101	-12.23 (37.06)	-43.84 (35.63)	195.58	189	6.58
Anxious	34	-37.71 (34.20)	-41.85 (42.24)	222.02	207	15.02
Bored	71	-53.77 (36.01)	-53.90 (42.85)	224.93	225	0.07
Stressed	41	-55.61 (35.08)	-36.29 (40.79)	236.87	243	6.13
Disappointed	55	-46.42 (43.21)	-23.67 (38.90)	242.98	261	18.02
Quadrant IV. Negative Valence x Positive Control/Power						
Frustrated	66	-39.11 (39.66)	-28.20 (46.53)	234.21	279	44.79
Envious	3	4.67 (59.14)	-4.33 (30.57)	132.88	297	164.12
Disgusted	9	-39.33 (56.50)	-18.56 (59.89)	244.74	315	70.26
Annoyed	15	-70.27 (35.79)	-58.00 (61.33)	230.46	333	102.54
Irritated	20	-21.25 (32.58)	23.20 (55.31)	317.51	351	33.49

The results highlight a wide range of differences between the expected and the observed disposition of each lexicalized emotion, going from almost absolute correspondence for *Bored* (0.22°) to more than a rotation of 90° for *Annoyed* (102.59°). The empirical position of some options is computed using only a few observations, as in the case of *Empathetic* (5), *Envious* (1), or *Disgusted* (5); whereas others show a good approximation with many observations, as it is the case for the aforementioned *Bored* (0.22° with 63 observations), *Amused* (7.20° with 94 observations), *Surprised* (5.72° with 35 observations), or *Stressed* (5.11° with 35 observations). The overall disposition of the observed affective space, though, corroborates the lack of orthogonality, highlighted earlier in the chapter, in the use of the two appraisal dimensions. In fact, the comparison between the graphical representations of the theoretical/expected circumplex in Figure 9.4 and the empirical/observed in Figure 9.5 confirms that most of the subjective feelings have been chosen with congruent ratings of *Valence* and *Control/Power*, that is when both appraisal dimensions are either positive or negative.

When the two appraisal dimensions are orthogonal, only *Surprised* and *Empathetic* in the bottom-right quadrant, and *Envious* and *Irritated* in the top-right quadrant – based on only a few observations, though – have been chosen with the expected combination of the appraisal dimensions. The other lexicalized emotions supposed to appear on the orthogonal combination of appraisals *moved* in another quadrant, depending on the *Valence* rating. That is, feelings from the Positive *Valence* x Negative *Control/Power* quadrant *moved* to the quadrant with both Positive appraisal dimensions; whereas feelings from the Negative *Valence* x Positive *Control/Power moved* to the quadrant with both Negative appraisal dimensions.

The phenomenon is consistent in both learning settings, with nevertheless some variations. Figure 9.6 shows the disposition of feelings with respect to the mean *Valence* and *Control/Power*, therefore in a Cartesian plane rather than in a circumplex. The choice of a different format, though, is only dictated by the need to simplify the display of the information, minimizing overlapping; there is thus no change in the underlying computational model. As Figure 9.6 shows, the overall tendency of co-variation in the two appraisal dimensions remains. Nevertheless, in the Synch./Collab. setting, there is more consistency in the bottom-right quadrant, the one characterized by Positive *Valence* x Negative *Control/Power*. Three out of five feelings appears in the expected quadrant, and a fourth one – *Relieved* – is close to the edge with the top-right quadrant.

Synthesis

The assessment of the link between the appraisal dimensions and the subjective feeling yielded mixed, but overall promising, results. On the one hand, the overall *accuracy* of the dynamic algorithm was around 0.8, which means that four out of five subjective feelings expressed by participants were included in the three buttons suggested on the interface. The accuracy was lower in the Asynch./Indiv. setting, though, which may suggest the need of a more nuanced expression in this condition.

On the other hand, data clearly confirm a major issue with the *Control/Power* appraisal dimension, which has a high correlation with the *Valence* dimension, a phenomenon already highlighted in the limitations of the Geneva Emotion Wheel in Section 4.2.3 and also observed in the usability test of the DEW illustrated in Section 4.3.4. The problem, though, does not seem to be unique to self-report tools. As already mentioned, Scherer and Fontaine (2018) encountered a similar difficulty with the use of the GRID instrument (Fontaine et al., 2013). As advocated by the authors, this problem requires future studies to find better solution.

9.4.4 Measures About the Perceived Usability of the Tool

Finally, the EAT will be appraised with respect to its usability, that is the perceived efficiency, effectiveness and satisfaction in using the tool. A usability measure, the System Usability Scale (SUS, Brooke, 1996), was administered in the empirical contribution of Chapter 8, but not of Chapter 8. The SUS was nevertheless administered in the usability test in Fritz (2015), which used the same configuration and task of the Synch./Collab.

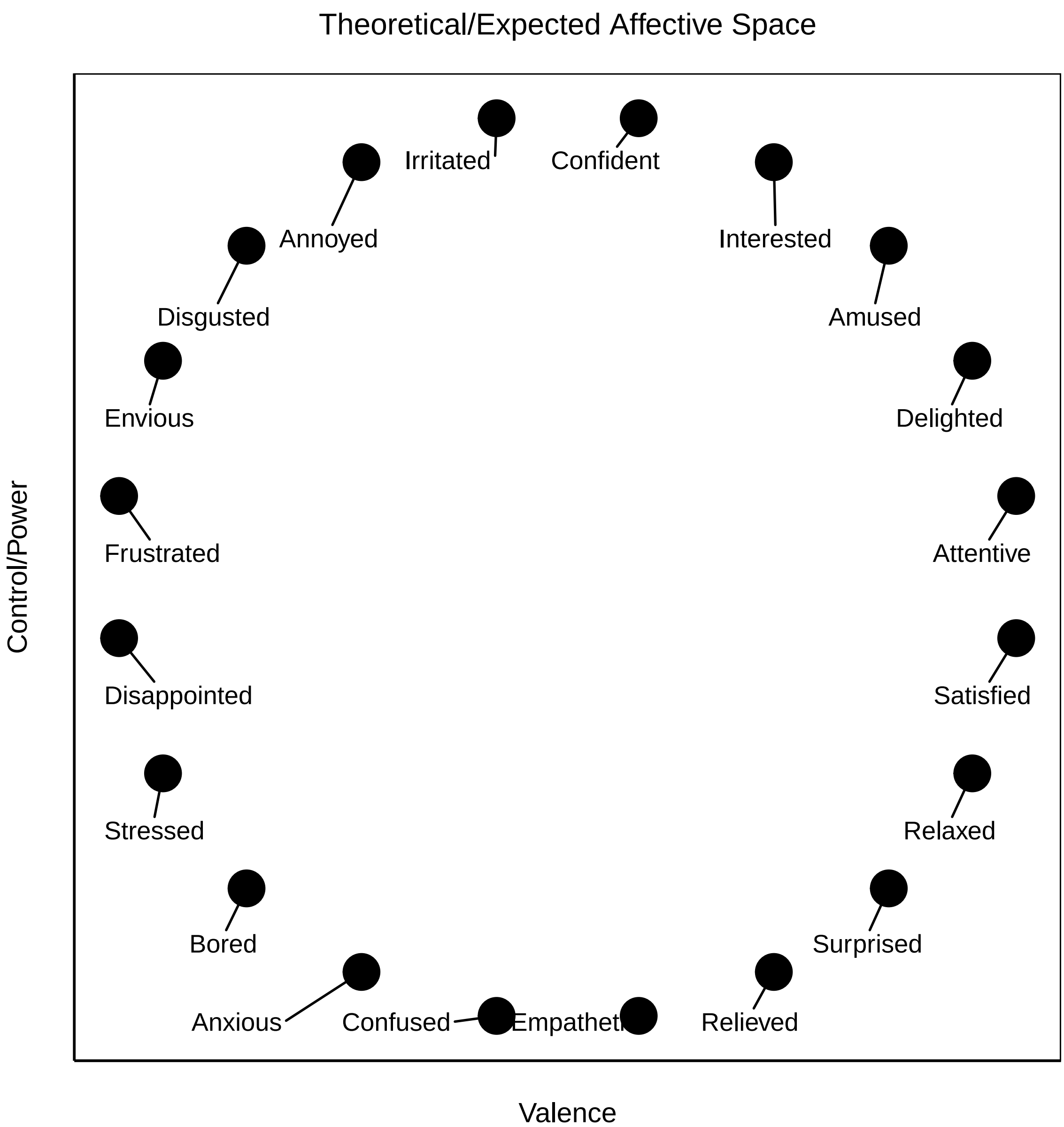


Figure 9.4: The theoretical/expected disposition of the EATMINT circumplex, reported from Figure 4.7 in Chapter 4.

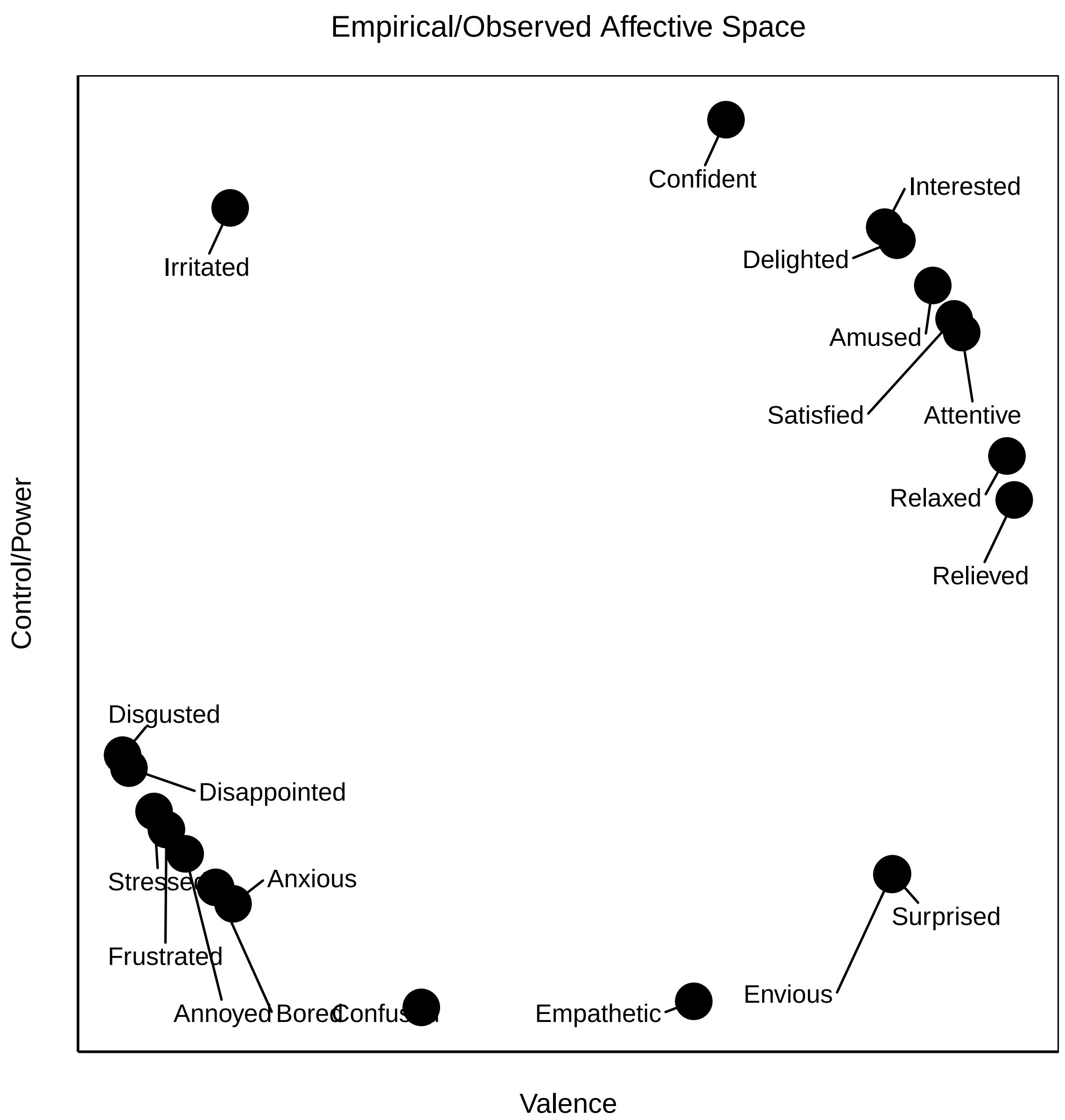


Figure 9.5: The empirical/observed disposition of the EATMINT's circumplex lexicalized emotions according to the actual average rate of participants.

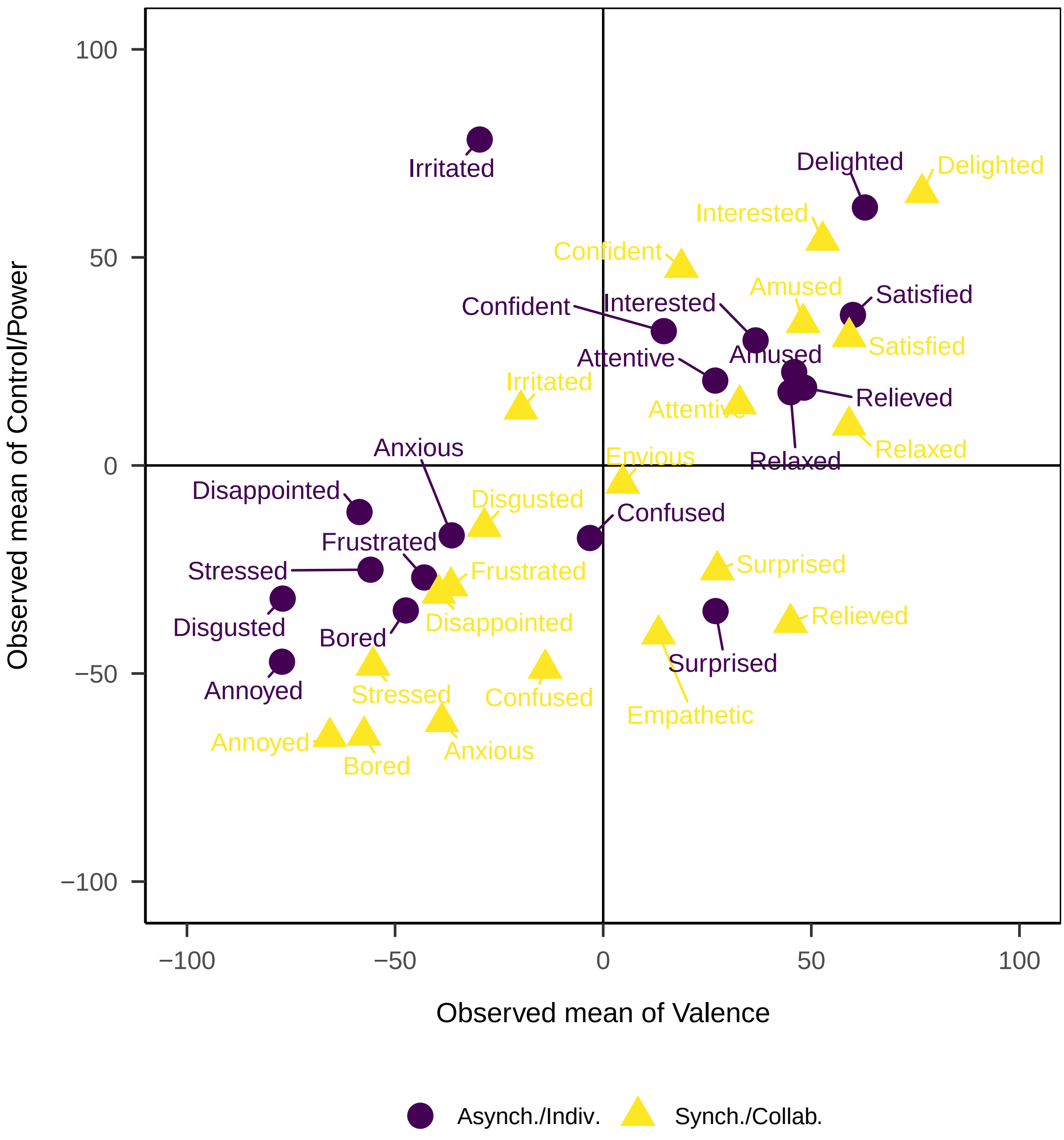


Figure 9.6: Comparing the empirical disposition of the two learning settings.

setting (but without the experimental conditions). The SUS scores of the usability test ($N = 16$) will therefore be used for the Synch./Collab. setting, alongside the scores obtained in the Asynch./Indiv. of Chapter 8 ($N = 26$).

The scale, using inverse rating for even items, allows to compute an overall score ranging from 0 (very poor perceived usability) to 100 (excellent perceived usability). Results of the SUS have been collected in the last decades in various published and unpublished reports. Thus, there is nowadays the possibility to better assess the overall score of the SUS, as well as of each of its ten items (Bangor et al., 2009; Lewis & Sauro, 2018; Sauro & Lewis, 2016).

Concerning the overall score of the SUS, Sauro and Lewis (2016) extrapolated a curved grading scale from 241 industrial usability studies and surveys. According to this scale, the average SUS overall score is $M = 68$. In the meantime, the same authors suggest that "it is becoming a common industrial goal to achieve a SUS of 80" (Lewis & Sauro, 2018, p. 161) as synonymous of a perceived good experience.

Table 9.8 shows the SUS score for the two learning settings, as well as the weighted overall mean. With an overall score of $M = 73.54$ ($SD = 12.83$), the EAT is perceived somehow halfway between the $M = 68$ empirical benchmark, and the target score of 80. As the table shows, the SUS score is consistent across the two learning settings, with a difference of less than 1 point.

Table 9.8: Score to the System Usability Scale (SUS, Brooke, 1996)

Condition	N	M	SD
Asynch./Indiv.	26	72.82	11.89
Synch./Collab.	14	74.88	14.81
Total	**40**	**73.54**	**12.83**

Furthermore, Lewis and Sauro (2018) collected data from 166 unpublished industrial usability studies or surveys, each one comprising a mean of the SUS overall score. The 166 means were computed from a total of 11'855 surveys. From these data, the authors retrieved benchmarks for each of the ten items of the SUS to reach the $M = 68$ empirical mean, or the target score of 80. As a reminder, the SUS items are the following:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Figure 9.7 depicts the score of each of the SUS items across settings. To ease the comparison, even items have already been reversed, so that for each item a higher score equals a higher perceived usability. The horizontal lines represent Lewis and Sauro (2018) benchmarks to reach the target score of 80. The original benchmarks refers to the 1-to-5 rating and have therefore been multiplied by a factor of 1.4 to map to the 1-to-7 scale used in both administrations of the SUS. Also the benchmarks have already been reversed for even items to ease the comparison.

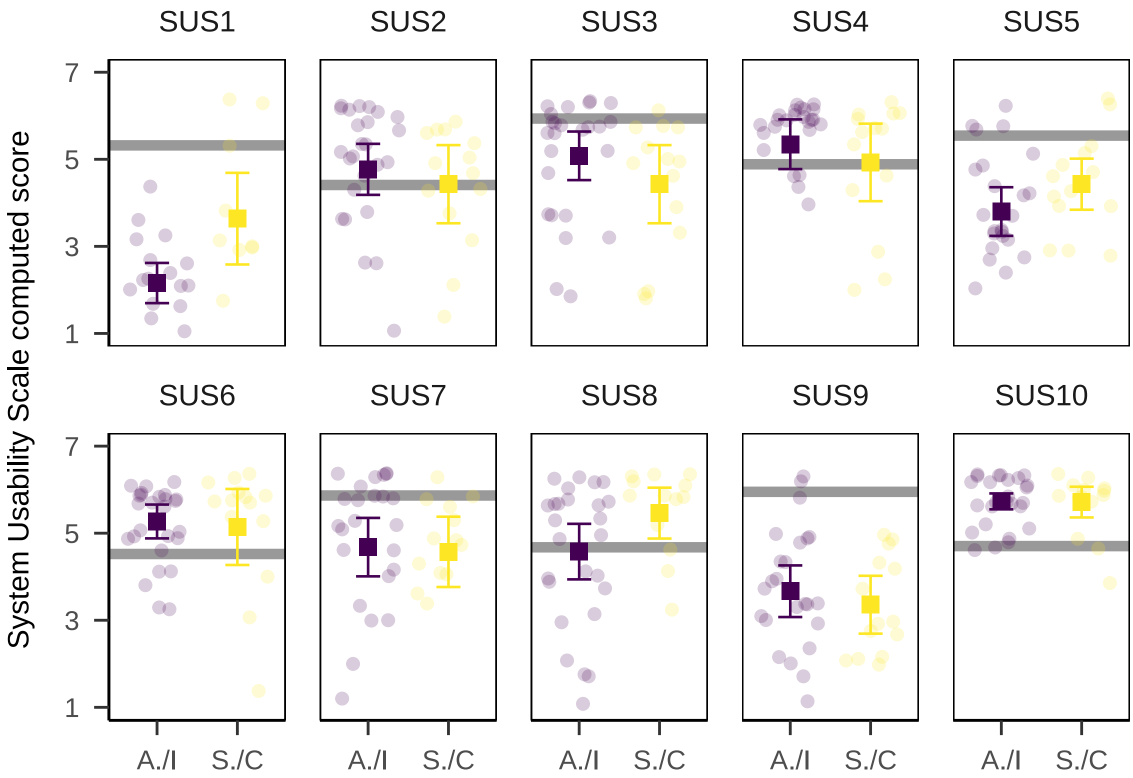


Figure 9.7: SUS scores on the single items, with the horizontal lines representing Lewis and Sauro (2018) benchmarks to reach the target score of 80, transformed to a 1-to-7 scale. Both items and benchmarks have already been reversed for even items. Bars represent 95% confidence intervals.

Data show substantial consistency across learning settings for most of the items, with differences in items SUS1 (frequency), SUS3 (ease of use), and SUS8 (intuition). The first item in particular is the one yielding the bigger discrepancy, with participants in the Synch./Collab. setting reporting a perspective use more frequent than the Asynch./Indiv. setting, which is consistent with the observed use, for instance in terms of number of emotions expressed. All in all, though, the EAT seems to possess an *intrinsic* perceived usability that holds – for good and for bad – in both settings.

Comparing the single items to Lewis and Sauro (2018) benchmarks highlights that half of the items are on-or-above the target, and half are below. The fact that all the odd items are below, and all even items are on-or-above the target may seem peculiar, but is consistent with the pattern of the benchmarks which are more demanding for odd items. Lewis and

Sauro (2018) do not mention anything specific about this pattern, but implicitly exclude it could be determined by the negative formulation of odd items, since previous findings collected by the same authors (Sauro & Lewis, 2011) seem to suggest that there is negligible impact of the negative formulation compared to a positive transformation of the odd items. An effect of the wording of items can nevertheless easily be tested or controlled by using an all-positive version of the SUS, which has already been adopted (*ibid.*).

According to the benchmarks, the EAT performs particularly bad in the frequency of use (SUS1) and in the confidence in the system (SUS9), which are two important dimensions of the scale. The integration of the different parts of the system (SUS5) also does not seem to convince learners, whereas ease of use (SUS3) and quickness of adoption (SUS7) remain below target, but less critically so. On the bright side, the EAT performs very well in learnability (SUS10), and well in consistency (SUS6). Furthermore, simplicity (SUS2), autonomy (SUS4), and intuition (SUS8) are aligned with the target score of 80.

A more contextual comparison of the usability of the EAT can be provided with Feidakis and colleagues (2014), who also used the SUS score to assess the usability of the *emot-control* (see the related works in Section 2.4). The overall SUS score obtained by the *emot-control* on $N = 29$ is $M = 67.81$ (SD not provided), thus around 6 points lower than the score observed for the EAT at hand. Contrary to the present results, though, learners in Feidakis and colleagues (*ibid.*) reported a higher score on the item about frequency (SUS1).

A preliminary synthesis highlights that the overall perception of the usability of the EAT is fairly good, especially considering the lack of previous experience with this type of device. On the other hand, there are also critical indications, such as the prospected frequency of use or the confidence in the system, which must be carefully considered.

9.5 Discussion

Overall, the analysis performed on the data at hand provided insightful cues upon which assess the EAT. With a dataset of more than one thousand expressed emotions, in particular, the emotion structure injected into the EAT have been put under the lens. The System Usability Scale (Brooke, 1996) also proved to be insightful, especially taking advantage of the benchmark and the direct comparison of a rating of another EAT (Feidakis et al., 2014). Finally, the two different conditions in which the EAT has been deployed fostered a preliminary assessment of its multi-purpose vocation of the toolbox. The discussion builds upon the four research questions for a thorough assessment.

9.5.1 Appraising the Appraisals

Data about the use of the appraisal dimensions draw a mixed picture about their efficacy. On the one hand, individually, each slider seem to be evaluated on its full spectrum, suggesting there is an interest in having them both. On the other hand, there is compelling evidence that the two dimensions are evaluated symmetrically, with an overall average correlation within respondents of $M_\rho = 0.47$, which even reached $M_\rho = 0.58$ in the Asynch./Indiv. setting. Even if it is not possible to deduce from a correlation which one is subordinate to the other, chances are that the *Valence* dimension has the upper hand, with the *Control/Power* dimension playing the sparring partner role. This would be consistent with other findings in fundamental emotion theory (see 4.2.3, but also in applied contexts. For instance, Lavoué et al. (2020) point out a problem with the *Control* dimension also stemming from a perspective informed by the Control-Value theory of achievement emotions (Pekrun, 2006; Pekrun & Perry, 2014). The authors explain:

> Regarding the perceived control of the situations in which students experience emotions, we first identify that they report mainly emotions associated with a low control, with a high percentage of negative emotions. This result is in line with studies that show that emotions associated with a low control provoke mainly negative emotions, such as anxiety and frustration (Pekrun 2006). Second, we also observe that a high number of situations are not associated with a specific level of control (high or low). We deduce that perceived control is an appraisal dimension that is not frequently used by students to explain their emotions, meaning that they may have difficulties in assessing their level of control over learning tasks.
> — Lavoué et al. (2020), p. 282

Except for the part about *negative* emotions being mainly provoked by low *Control* (this does not hold in the Component Process Model where, for instance, *Frustration* is characterized by high *Control/Power*), the reminder of the description seems to perfectly fit the data at hand. The case of the single student expressing more than 60 emotions without touching the *Control/Power* dimension is in this sense revealing.

The problem with the rating of the *Control/Power* dimension was also one of the main reasons to provide researchers and practitioners with the possibility to configure in the toolbox the underlying affective space and the way each appraisal dimensions is prompted on the interface. The scaling of the computational model to an n-dimensional structure was also influenced by empirical results, alongside the theoretical considerations illustrated in the presentation of the model in chapter 5.

Getting rid of the *Control/Power* dimension altogether seems nevertheless excessive. Depending on the theoretical approach adopted, this appraisal/dimension may play a prominent role in emotion elicitation and differentiation (Fontaine et al., 2021; Pekrun & Perry, 2014; Scherer, 2009b; Scherer et al., 2018). Rewording of the way the rating is prompted may therefore be a solution, even if in this case it hasn't apparently led to any substantial improvement. A second options consists in training: explaining beforehand to learners/participants what is *really* meant by this dimension, with examples and prototypical vignettes.

A third options is more revolutionary and provoking. Given evidence that *Control/Power* is subordinate to the *Valence* dimension (Scherer & Fontaine, 2018), rather than getting rid of the *Control/Power* dimension, one can attempt to bypass the *macro Valence* instead (Erbas et al., 2015; Shuman et al., 2013). Not asking respondents to rate whether the situation is pleasant or unpleasant may free their cognitive evaluation of a looming presence. As suggested by Erbas et al. (2015), focusing on *Valence* alone seem to impoverish the overall gamut of potential emotional experience. Especially in an overall context in which self-reflection may be of pivotal importance, proposing more *nuanced* appraisals dimensions can also have benefits in widening learners perspective on their emotional experience. This would not only impact the analysis of the causes of how a learner is feeling, but possibly also empower the strategies for emotion regulation. If one is focused exclusively on whether the situation is *positive* or *negative* with respect to *Valence*, then she may be tempted to apply a linear way of action: lean towards the *positive,* full stop. Reflecting on the emotional experience from other points of view may highlight the functional role of emotion in general, not only positively valenced emotions.

To sum up, the potential of the appraisal dimensions seem for the moment quite unexpressed. Whether this is due to contingent factors or to a deeper structural reason remains an open question of pivotal importance for the overall assumption defended by the thesis. Appraisals rest at the core of an instrumental use of emotional awareness in computer-mediated learning environments, and the assessment of their contribution is therefore a necessary condition for a severe test of the EAT. Leveraging on the flexibility in configuring instances of the toolbox, a next step in the assessment could be to experimentally compare two versions of the EAT *powered* by different appraisal dimensions.

9.5.2 EATMINT's Lexicalized Emotions Are Representative of Learners' Experience

At first blush, the lexicalized emotions proposed by the EATMINT circumplex (see Section 4.3.3) seem to provide learners with adequate options to represent their *true* underlying subjective feeling. Around 90 percent of the expressed emotions adopted one of the lexicalized emotion from the pre-defined list in the affective space.

A distinction between the two settings nevertheless has also emerged, quantified in an decreased accuracy of more than 10 percentage points in the Asynch./Indiv. setting. This difference leans toward the hypothesis that in longitudinal and non-collaborative settings, learners may need a wider and more *nuanced* emotional vocabulary to express how they feel. This may be due to the fact that in such condition, the *object focus* (Moors, 2009; Pekrun & Perry, 2014; Sander, 2013) of the emotional experience can be wider compared to a synchronous and collaborative setting, where the domain of events may be more tightly bound to what is happening *hic et nunc*. As stated by Pekrun in the Control-Value theory (Pekrun & Perry, 2014), retrospective and prospective emotions can also play a prominent role in emotion related to learning. Learners may project more stable characteristics of themselves, bestowing to the emotional experience a more complex and overarching meaning with respect to the immediacy and volatility of synchronous exchanges (Rimé, 2005, 2009). These considerations are nevertheless at the stage of speculation given the limitations of the sample at hand.

A more pragmatic discussion may relate to the number of lexicalized emotions to propose in the pre-defined list. On the one hand, there are contributions highlighting a restricted domain of emotion term that seem to account for the most frequent emotions experienced in computer-mediated learning environments (D'Mello & Graesser, 2012; Gaëlle Molinari et al., 2016; Reis et al., 2015). On the other, emotion differentiation play an important role in representing the nuances of how a person feels, to herself or to others (Erbas et al., 2015; Grandjean et al., 2008; Torre & Lieberman, 2018). The data at hand strike a balance between the two: it is evident that there are lexicalized emotions that haven been chose consistently more often than others, but at the same time each option has been chosen at least by some participants. *Amused*, *Attentive*, *Interested*, *Confused* and *Bored* were among the most frequently chosen ones, whereas *Envious*, *Disgusted*, *Empathetic*, *Annoyed*, and *Relaxed* the least frequent ones. There were also differences between settings, with some lexicalized emotions being chosen more often in one setting than the other, like *Amused*, *Relieved*, *Confused*, *Surprised*, and *Satisfied*. This seems to confirm that the context has an important role in determining whether there are lexicalized emotion who are more or less *prototypical* according to the situation at hand (Gentsch et al., 2017).

It is therefore impossible to figure out whether there are a number and kind of lexicalized emotions that can fit any situation, even though the EATMINT circumplex seems to adapt at least to the settings at hand. When a number of closed options do not have a particular importance – for instance when there is some form of intelligence that must respond to a limited domain of inputs, see Section **??** – it is argued here that it would be better to provide learners with a sufficient number of lexicalized emotions to choose from. This may once again broaden their perspective on emotion differentiation and empower their ability to reflect on their own emotional experience, as well as increasing the chances to convey a strategic signal to others (Scherer, 2007). Since all the contributions adopted a list of 20 lexicalized emotions, it would be easy to confirm that this is a sufficient number. But obviously this is a superficial assessment. Once again, if this information is crucial, it may be more accurately determined empirically, comparing for instance affective spaces that vary on the number of pre-defined options.

A last point concerning the EATMINT circumplex may be cited with respect to the *Zeitgeist*. The affective space use a list of adjectives rather than nouns for the lexicalized emotions. In French, this has the shortcoming of providing only adjective with the male declination at the end. To the best of my knowledge, there is no compelling evidence that the distinction between an adjective and a noun can influence the subjective feeling. If this is the case, then it would be more appropriate to switch the EATMINT circumplex on a noun-based format, which would be more inclusive. If avoiding the noun format is of essence, a more neutral form such as *intéressant* (interesting), rather than *intéressé-e* (interested) could be adopted instead.

To sum up, the number and kind of predefined lexicalized emotion may play a prominent role in sustaining learners in expressing how they feel. The options provided by the EATMINT circumplex seem to fit this purpose, adapting to different participants and environments, even if differences between the Asynch./Indiv. and Synch./Collab. settings could be observed. Given that the same circumplex has been used throughout the contributions, there is nevertheless no counter-factual evidence to better assess this claim. In this sense, a very severe test would be to adopt a *pure* constructivist approach (Barrett, 2006, 2017, 2018) and force learners to type in an emotion term, without providing any predefined option. It would be interesting to compare how many lexicalized emotions of the EATMINT circumplex would spontaneously be chosen as representative of learners subjective feelings.

9.5.3 The Computational Model Is a Good Enough Heuristics Most of the Time

The assessment of the parsimonious computational model presented in Chapter 5 provided mixed evidence about its interest and accuracy. The interest seems warranted, since participants found a representative subjective feeling in the three options proposed as buttons around 3 out of 4 times, with a proportion nearing 4 out of 5 in the Synch./Collab. setting. Considering that the system proposed only 3 out of 20 lexicalized emotion, the performance seems to corroborate that is indeed possible to predict the emotional experience of a person based on how she has appraised the situation (Fontaine et al., 2021; Gentsch et al., 2017; Scherer & Fontaine, 2018; Scherer & Meuleman, 2013; Scherer et al., 2018).

The efficacy of the computational model nevertheless depends from the interplay between the accuracy of the underlying affective space on the one hand (Fontaine et al., 2021; Gillioz et al., 2016; Scherer et al., 2006; Scherer, Shuman, et al., 2013), and the respondents' *adequate* interpretation of the appraisal criteria on the other hand (Scherer, 2007, 2021a).

With the data at disposal it is evident that the *observed* affective space, retrieved from participants' actual rating of the appraisal dimensions when a given subjective feeling was expressed, does not overlap with the theoretically-driven disposition in the EATMINT circumplex. Given that the disposition of the lexicalized emotions in the EATMINT circumplex has not been validated, the possibility of a problem with the underlying affective space cannot be formally ruled out. Nonetheless, it is safe to assume that the problem stems, at least principally, from the evaluation of the appraisal dimensions, as already discussed above. When the *Valence* and *Control/Power* are rated symmetrically, it is inevitable that all subjective feelings align in the upper-right and bottom-left quadrants of a two-dimensional space.

Given its computational parsimony and the benefit in reducing the number of concurrent discrete emotions on the interface, the computational model can be at least retained as a good enough heuristics from a usability stand point. It allows respondents to perform a first skim, which is sufficient most of the time, but does not bind the choice to the pre-defined options. In this regard, the use of an alternative way to input a subjective feeling outside the proposed options is not only a good practice in emotion self-report (Mortillaro & Mehu, 2015; Scherer, 2005), but becomes essential to the possibility to express a strategic signal (Scherer, 2007).

To sum up, the computational model remains a promising feature of the EAT as long as it does not have the pretension to accurately predict the emotional experience of respondents every time and in every situation. For a more severe assessment of the interest and accuracy of the computational model, though, a direct comparison with an alternative mechanism should be implemented. In this regard, the toolbox proposes a random algorithm that can be applied to any underlying affective space. A direct comparison between the computational model and a random sub-setting may help to quantify more specifically the benefit in endowing the EAT with a probabilistic link between appraisal dimensions and subjective feelings.

9.5.4 Usability Is Not Bad – But It Is Not Enough

The last assessment takes a more subjective perspective and, rather then the concrete use of the EAT, focuses on the perception of its usability. The System Usability Scale (Brooke, 1996) was administered in the usability test and at the end of the longitudinal design of Chapter 8. The overall rating of the tool was of $M = 73.5$, that is, halfway between the benchmarks provided by Lewis & Sauro (2018): the overall average of 68 points and the target of 80 points for good usability. In a more direct comparison with the rating of the *emot-control* tool in Feidakis et al. (2014), for which an average of 67.8 was observed, the EAT therefore received a rating of around 6 points higher. The rating on the SUS can therefore be retained as promising, even if some dimensions (frequency of use especially) scored far from the target benchmark to reach 80 points.

More than the score in itself, though, the overall assessment of the perception of the tool should be ameliorated by taking into account a wider perspective. In fact, the SUS is a very reliable and valid scale, adopted in different contexts. Nevertheless, as the name implies, it focuses on usability alone, whereas there is growing consensus in trying to assess the user experience more globally (Lallemand & Gronier, 2017; Law, Van Schaik, & Roto, 2014; Tullis & Albert, 2013a). Tullis & Albert (2013a) draws in this sense a useful distinction:

> Usability is usually considered the ability of the user to use the thing to carry out a task successfully, whereas user experience takes a broader view, looking at the individual's entire interaction with the thing, as well as the thoughts, feelings, and perceptions that result from that interaction.
> — Tullis & Albert (2013a), p. 5

Given the importance that is attributed to the affective experience by the EAT itself, it would be coherent to follow through with a more holistic assessment, integrating usability measures with scales such as the AttrakDiff (Hassenzahl, Burmester, & Koller, 2003), the UEQ (Laugwitz, Held, & Schrepp, 2008), or the meCUE (Minge, Thüring, Wagner, & Kuhr, 2017). These scales provide a more thorough assessment of a technological device, considering for instance also the look-and-feel of the tool, which is known to not only impact the perception of the tool, but also its use (Lallemand & Gronier, 2017; Tullis & Albert, 2013a; Vermeeren, Law, Roto, Obrist, & Väänänen-Vainio-Mattila, 2010).

In the meantime, usability and user experience would cover only the perspective of the individual learner that interact with the tool. A more thorough assessment of the perception of the EAT should therefore also be guided by a more comprehensive theoretical framework in human-computer interaction. One potential candidate is the widely adopted Technology Acceptance Model [TAM; Venkatesh, Morris, Davis, & Davis (2003); F. D. Davis (1989); F. D. Davis & Venkatesh (1996); Venkatesh & Davis (2000)]. The main assumptions of the model state that the actual use of a technological device is determined in particular by two intervening factors: perceived usefulness and perceived ease of use. The TAM may therefore be integrated alongside a more specific scale such as the tentative Emotion Awareness Usefulness (EAU) scale introduced in Chapter 8, which targets specifically the use of an EAT.

Another advantage of introducing a more overarching framework in human-computer interaction concerns the perspective of practitioners who may be interested in endowing their courses with an EAT. A meta-analysis conducted by Scherer*, Siddiq, & Tondeur (2019)[1] illustrates how teacher's acceptance of a technology is an integral part in determining to what extent it can be fruitfully adopted by students. For instance, the RULER approach (Brackett et al., 2019; Brackett et al., 2012a; Hoffmann et al., 2020), of which the Mood Meter application is an integral part, proposes training for its use in the classroom.

To sum up, the assessment of the perception about the use of the EAT has so far focused on the usability perspective, which was pertinent during the first phases in the interaction design process. Even though the response has been promising so far, as the iterations progress, the assessment should widen and take into consideration a more holistic perception of the EAT. One way to move forward is by integrating user experience scales, and also an overarching theoretical framework such as the TAM, which may also implicate the perception of practitioners interested in endowing their classes with the use of an EAT.

9.6 Conclusion

This chapter proposed an overall data-driven assessment of the use and perception of the EAT, as well as a comparison between the use in an asynchronous and individual environment, versus a synchronous and collaborative one. The primary purpose of the study was to determine to what extent the EAT meets learners' need, in particular by evaluating whether learners' take full advantage of the emotional structure implemented into the tool in expressing their emotional states. Results of the overall and comparative assessment provide mixed evidence.

On the one hand, the tool and the underlying EATMINT affective space seem to adapt fairly well to different settings. For instance, the 20 lexicalized emotions proposed in the circumplex seem to be sufficient to convey most of learners' emotional experiences. Furthermore, the algorithm linking the two appraisal dimensions with the proposed subjective feelings provided a good heuristic, consistent across different settings. These indicators suggest that the EAT possesses a sort of *intrinsic* value, which may be adaptable to different contexts. This does not mean, though, that the EAT will be perceived as useful regardless of other determinants, such as the task at hand or the overall instructional design. For instance, the comparison suggests that learners' may need a more nuanced emotional expression in an asynchronous and individual setting.

On the other hand, results also corroborate and extend some problems already emerged both theoretically and in applied contexts, such as the lack of orthogonality of the two appraisal dimensions *Valence* and *Control/Power*. This kind of issues, though, does not seem limited to technical elements of the EAT, but rather extend to more fundamental aspects of emotional awareness, which will be considered more thoroughly in the general discussion in Part IV of the thesis. This chapter concludes with the limitations and the contribution of this study in particular.

[1] I took the liberty to add to Ronny Scherer last name an * at the end, so that the distinction from Klaus R. Scherer will not entail – following the orthodoxy of APA rules – the full name stated in the many contributions where Klaus R. Scherer is involved. I hope the authors, if they will ever read the manuscript, would not mind.

9.6.1 Limitations

The study proposed a mixture between a secondary data analysis (Weston et al., 2019) and a small internal meta-analysis of one's own studies (Goh et al., 2016). By staying in between the two, the risk is that neither of the qualities of both methods has been truly exploited. The sample size at hand remain in fact relatively small, and the research questions were not clearly stated beforehand, but emerged as long as data were collected or analyzed for the two empirical contributions of the thesis.

Even the sample size of more than one thousand expressed emotions, which seems huge, must be put into perspective. In a provocative comparison, it may be stated that with two ratings between -100 and 100, as well as 20 lexicalized emotions as possible outcome (not taking into account custom responses), there is a total of 808'020 possible combinations between the three measures that composes an observation in the three datasets at hand. A greater number of emotion expressed would also provide the possibility to fit more complex models on the data at hand (Guest & Martin, 2021; Marsella et al., 2010b; McElreath, 2020; Rodgers, 2010).

9.6.2 Study's Overall Contribution

In spite of the limitations, the study provided a first attempt in comparing measures obtained through the same toolbox, but with different instances of the EAT. Even though the instances in this specific case were not very different, this is still a first example of the kind of analyses that would be possible when data is produced and shared in a common format (Gilmore et al., 2018; Levenstein & Lyle, 2018).

From a date-driven perspective, the study explored topics that have both practical and theoretical relevance. For instance, the analyses allowed to investigate the rating on appraisal dimensions, revealing a non-independence between the *Valence* and the *Control/Power* appraisals also encountered in fundamental and applied research (Erbas et al., 2015; Lavoué et al., 2020; Scherer & Fontaine, 2018; Shuman et al., 2013). Through the inner functioning of the EAT, it was also possible to compare a theoretically driven and an *empirically rated* affective spaces. In the case of lack of resources for a more thorough validation of an affective space – such as using the GRID instrument (Scherer, Fontaine, et al., 2013; Scherer, Shuman, et al., 2013) – this method may represent a viable approximation to validate or, at least, assess the accuracy of an affective space. In this regard, from an Open Science perspective, it may be considered to provide some common analyses that can be performed through data exported from the toolbox in an automated process. A good candidate for this role would be a package in the R programming language environment.

Part IV

Concluding Remarks

Chapter 10

General Discussion

Conclusion

Appendix

References

Adolphs, R., & Anderson, D. J. (2018). *The neuroscience of emotion: A new synthesis*. Princeton: Princeton University Press.

Adolphs, R., & Andler, D. (2018). Investigating Emotions as Functional States Distinct From Feelings. *Emotion Review*, *10*(3), 191–201. http://doi.org/10.1177/1754073918765662

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, *74*, 187–195. http://doi.org/10.1016/j.jesp.2017.09.004

Andriessen, J., Baker, M., & der Puil, C. V. (2011). Socio-cognitive tension in collaborative working relations. In S. Ludvigsen, A. Lund, I. Rasmussen, & R. Saljo (Eds.), *Learning across sites: New tools, infrastructures and practices*. London, UK: Pergamon.

Anvari, F., & Lakens, D. (2019). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology*, 1–21. http://doi.org/10.1080/23743603.2019.1684822

Arguedas, M., Daradoumis, T., & Xhafa, F. (2016). Analyzing how emotion awareness influences students' motivation, engagement, self-regulation and learning outcome. *Educational Technology and Society*, *19*(2), 87–103.

Arguedas, M., Xhafa, F., Casillas, L., Daradoumis, T., PeÃÂa, A., & CaballÃÂ, S. (2018). A model for providing emotion awareness and feedback using fuzzy logic in online learning. *Soft Computing*, *22*(3), 963–977. http://doi.org/10.1007/s00500-016-2399-0

Arguedas, M., Xhafa, F., Daradoumis, T., & Caballe, S. (2015). An Ontology about Emotion Awareness and Affective Feedback in Elearning. In *Proceedings - 2015 International Conference on Intelligent Networking and Collaborative Systems, IEEE INCoS 2015* (pp. 156–163). http://doi.org/10.1109/INCoS.2015.78

Armony, J., & Vuilleumier, P. (Eds.). (2013). *The Cambridge handbook of human affective neuroscience*. Cambridge ; New York: Cambridge University Press.

Arnold, M. B. (1960). *Emotion and personality psychological aspects* (Vol. 1). New York, NY: Columbia University Press.

Avry, S. (2021). Explicit Sharing of Emotions Improves the Relationship of Groups with Lower Dispositions to Regulate Emotions in Collaborative Problem-solving, 4.

Avry, S., Chanel, G., Bétrancourt, M., & Molinari, G. (2020). Achievement appraisals, emotions and socio-cognitive processes: How they interplay in collaborative problem-solving? *Computers in Human Behavior*, *107*, 106267. http://doi.org/10.1016/j.chb.2020.106267

Avry, S., Molinari, G., Bétrancourt, M., & Chanel, G. (2020). Sharing Emotions Contributes to Regulating Collaborative Intentions in Group Problem-Solving. *Frontiers* *in Psychology*, *11*, 1160. http://doi.org/10.3389/fpsyg.2020.01160

Bahreini, K., Nadolski, R., & Westera, W. (2016). Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, *24*(3), 590–605. http://doi.org/10.1080/10494820.2014.908927

Baker, M., Andriessen, J., & Järvelä, S. (Eds.). (2013). *Affective Learning Together. Social and emotional dimensions of collaborative learning.* Oxon, UK: Routledge.

Baker, R. S. J. D., D'Mello, S., Rodrigo, Ma. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human Computer Studies*, *68*(4), 223–241. http://doi.org/10.1016/j.ijhcs.2009.12.003

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., Laganke, J. S., & Warren, W. (2002). Computer-Mediated Communication and Group Decision Making: A Meta-Analysis. *Organizational Behavior and Human Decision Processes*, *87*(1), 156–179. http://doi.org/10.1006

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, *4*(3), 114–123. http://doi.org/66.39.39.113

Bänziger, T., Grandjean, D., & Scherer, K. R. (2009). Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*, *9*(5), 691–704. http://doi.org/10.1037/a0017088

Baralou, E., & McInnes, P. (2013). Emotions and the spatialisation of social relations in text-based computer-mediated communication. *New Technology, Work and Employment*, *28*(2), 160–175. http://doi.org/10.1111/ntwe.12012

Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *10*(1), 20–46. http://doi.org/10.1207/s15327957pspr1001_2

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. http://doi.org/10.1093/scan/nsw154

Barrett, L. F. (2018). *How emotions are made: The secret life of the brain* (Paperback edition). London: PAN Books.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. http://doi.org/10.1177/1529100619832930

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, *15*(6), 713–724. http://doi.org/10.1080/02699930143000239

Barsade, S. G. (2002). The ripple effects: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, *47*(47), 644–75.

Barsade, S. G., & Knight, A. P. (2015). Group Affect. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 21–46. http://doi.org/10.1146/annurev-orgpsych-032414-111316

Bates, A. W., & Bates, T. (2005). *Technology, e-learning and distance education.* Psychology Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018, May). Parsimonious Mixed Models. Retrieved from `https://arxiv.org/abs/1506.04967`

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014, June). Fitting Linear Mixed-Effects Models using Lme4. Retrieved from `https://arxiv.org/abs/1406.5823`

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition*, *55*(1), 30–40. http://doi.org/10.1016/j.bandc.2003.04.001

Beege, M., Schneider, S., Nebel, S., Hässler, A., & Günter, D. R. (2018). Mood-affect congruency. Exploring the relation between learnersâ€™ mood and the affective charge of educational videos. *Computers & Education*, *123*, 85–96. http://doi.org/10.1016/j.compedu.2018.05.001

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. http://doi.org/10.1038/s41562-017-0189-z

Berset, P. (2018). *Visualisation des données de recherche concernant le partage des émotions : Une approche centrée utilisateurs* (Master Thesis). Geneva University, Geneva, Switzerland.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2017). Visualization of Eye Tracking Data: A Taxonomy and Survey: Visualization of Eye Tracking Data. *Computer Graphics Forum*, *36*(8), 260–284. http://doi.org/10.1111/cgf.13079

Bloom, P. (2016). Empathy and Its Discontents. *Trends in Cognitive Sciences*, *xx*, 1–8. http://doi.org/10.1016/j.tics.2016.11.004

Blunden, H., & Brodsky, A. (2020). Beyond the Emoticon: Are There Unintentional Cues of Emotion in Email? *Personality and Social Psychology Bulletin*, 0146167220936054. http://doi.org/10.1177/0146167220936054

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, *6*. http://doi.org/10.3389/fpubh.2018.00149

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043–1045. http://doi.org/10.1016/j.chb.2010.07.014

Boden, M. T., & Thompson, R. J. (2015). Facets of emotional awareness and associations with emotion regulation and depression. *Emotion*, *15*(3), 399–410. http://doi.org/10.1037/emo0000057

Boehner, K., DePaula, R., Dourish, P., & Sengers, P. (2007). How emotion is made and measured. *International Journal of Human-Computer Studies*, *65*(4), 275–291. http://doi.org/10.1016/j.ijhcs.2006.11.016

Bonanno, G. A., Papa, A., Lalande, K., Westphal, M., & Coifman, K. (2004). The importance of being flexible: The ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science*, *15*(7), 482–487. http://doi.org/10.1111/j.0956-7976.2004.00705.x

Borenstein, M. (Ed.). (2009). *Introduction to meta-analysis.* Chichester, U.K: John Wiley & Sons.

Brackett, M. A. (2019). *Permission to feel: Unlocking the power of emotions to help our kids, ourselves, and our society thrive.*

Brackett, M. A., Bailey, C. S., Hoffmann, J. D., & Simmons, D. N. (2019). RULER: A Theory-Driven, Systemic Approach to Social, Emotional, and Academic Learning. *Educational Psychologist*, *54*(3), 144–161. http://doi.org/10.1080/00461520.2019.1614447

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012b). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*(2), 218–224. http://doi.org/10.1016/j.lindif.2010.10.002

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012a). Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum. *Learning and Individual Differences*, *22*(2), 218–224. http://doi.org/10.1016/j.lindif.2010.10.002

Bradley, M., & Lang, P. J. (1994). Measuring Emotion : The Self-Assessment Semantic Differential Manikin and the. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(I), 49–59.

Broekens, J., & Brinkman, W. P. (2013). AffectButton: A method for reliable and valid affective self-report. *International Journal of Human Computer Studies*, *71*(6), 641–667. http://doi.org/10.1016/j.ijhcs.2013.02.003

Brooke, J. (1996). SUS: A quick and dirty usability scale. In B. A. Weerdmester & I. L. McClelland (Eds.), *Usability evaluation in industry.* London, UK: Taylor & Francis.

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition & Emotion*, *24*(3), 377–400. http://doi.org/10.1080/02699930902975754

Brosch, T., Pourtois, G., Sander, D., & Vuilleumier, P. (2011). Additive effects of emotional, endogenous, and exogenous attention: Behavioral and electrophysiological evidence. *Neuropsychologia*, *49*(7), 1779–1787. http://doi.org/10.1016/j.neuropsychologia.2011.02.056

Brosch, T., & Sander, D. (2013). The Appraising Brain: Towards a Neuro-Cognitive Model of Appraisal Processes in Emotion. *Emotion Review*, *5*(2), 163–168. http://doi.org/10.1177/1754073912468298

Brosch, T., Scherer, K. R., Grandjean, D., & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, *143*, 1–10. http://doi.org/10.4414/smw.2013.13786

Brown, V. A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920960351. http://doi.org/10.1177/2515245920960351

Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior*, *27*(3), 1114–1117. http://doi.org/10.1016/j.chb.2010.07.012

Buder, J., Bodemer, D., & Ogata, H. (2021). Group Awareness. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (pp. 295–313). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-030-65291-3_16

Caicedo, D., & van Beuzekom, M. (2006). *How do you feel? An assessment of existing tools for the measurement of emotions and their application in consumer product research.* Delft University of Technology, Department of Industrial Design.

Calvo, R. A., & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *T. Affective Computing*, *1*, 18–37. http://doi.org/10.1109/T-AFFC.2010.1

Calvo, R. A., & D'Mello, S. (2012). Frontiers of Affect-Aware Learning Technologies. *IEEE Intelligent Systems*, *27*(6), 86–89. http://doi.org/10.1109/MIS.2012.110

Calvo, R. A., D'Mello, S., Gratch, J., Kappas, A., & Graesser, A. C. (2015). Feeling, Thinking, and Computing with Affect-Aware Learning Technologies. In R. A. Calvo, S. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford Handbook of Affective Computing.* Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199942237.013.032

Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, *4*(2), 75–100.

http://doi.org/10.1037/1076-898X.4.2.75

Cassidy, S. A., Dimova, R., Giguère, B., Spence, J. R., & Stanley, D. J. (2019). Failing Grade: 89. *Advances in Methods and Practices in Psychological Science*, *2*(3), 233–239. http://doi.org/10.1177/2515245919858072

Cereghetti, D., Molinari, G., Chanel, G., Pun, T., & Bétrancourt, M. (2015). Sharing emotions during a computer-mediated collaborative task: A dual eye-tracking study. In *EARLI 2015, European Conference for Research on Learning and Instruction.* Limassol, Cyprus.

Cernea, D., & Kerren, A. (2015). A survey of technologies on the rise for emotion-enhanced interaction. *Journal of Visual Languages & Computing*, *31*, 70–86. http://doi.org/10.1016/j.jvlc.2015.10.001

Cernea, D., Weber, C., Kerren, A., & Ebert, A. (2014). Group Affective Tone Awareness and Regulation through Virtual Agents. In *IVA 2014 Workshop on Affective Agents* (pp. 9–16). Boston, MA.

Chambers, C. (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice.* Princeton: Princeton University Press.

Chanel, G., Lalanne, D., Lavoué, E., Lund, K., Molinari, G., Ringeval, F., & Weinberger, A. (2016). Grand Challenge Problem 2: Adaptive Awareness for Social Regulation of Emotions in Online Collaborative Learning Environments. In *Grand Challenge Problems in Technology-Enhanced Learning II: MOOCs and Beyond* (pp. 13–16). Cham. http://doi.org/10.1007/978-3-319-12562-6

Chanel, G., Molinari, G., Cereghetti, D., Pun, T., & Bétrancourt, M. (2013). Assessment of Computer-Supported Collaborative Processes using Interpersonal Physiological and Eye-Movement Coupling. In *ACII 2013, 5th Int. Conf. Affective Computing and Intelligent Interaction* (pp. 2–5). Geneva, Switzerland.

Cherniss, C. (2010). Emotional Intelligence: Toward Clarification of a Concept. *Industrial and Organizational Psychology*, *3*(2), 110–126. http://doi.org/10.1111/j.1754-9434.2010.01222.x

Cheshin, A., Rafaeli, A., & Bos, N. (2011). Anger and happiness in virtual teams: Emotional influences of text and behavior on others' affect in the absence of non-verbal cues. *Organizational Behavior and Human Decision Processes*, *116*(1), 2–16. http://doi.org/10.1016/j.obhdp.2011.06.002

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (Vol. 13, pp. 127–149). http://doi.org/10.1037/10096-006

Colombetti, G. (2009). Appraising valence. *Journal of Consciousness Studies*, *12*(8-10), 103–126. http://doi.org/10.1016/j.amjmed.2012.04.013

Colombetti, G. (2014). *The feeling body: Affective science meets the enactive mind* (First

MIT Press paperback edition). Cambridge,MA: The MIT Press.

Cooper, A., Reinman, R., & Cronin, D. (2007). *About Face 3: The Essential of Interaction Design.* Indianapolis, IN: Wiley Publishing Inc.

Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, *114*(38), E7900–E7909. http://doi.org/10.1073/pnas.1702247114

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. C. (2008). Emote aloud during learning with AutoTutor: Applying the Facial Action Coding System to Cognitive–Affective states during learning. *Cognition & Emotion*, *22*(5), 777–788. http://doi.org/10.1080/02699930701516759

Cress, U., Rosé, C., Wise, A. F., & Oshima, J. (Eds.). (2021). *International Handbook of Computer-Supported Collaborative Learning.* Cham: Springer International Publishing. http://doi.org/10.1007/978-3-030-65291-3

Csikszentmihalyi, M., & Larson, R. (2014). Validity and Reliability of the Experience-Sampling Method. In *Flow and the Foundations of Positive Psychology* (pp. 35–54). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-017-9088-8_3

D'Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, *105*(4), 1082–1099. http://doi.org/10.1037/a0032674

D'Mello, S., & Graesser, A. C. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. http://doi.org/10.1016/j.learninstruc.2011.10.001

D'Mello, S., & Graesser, A. C. (2015). Feeling, Thinking, and Computing with Affect-Aware Learning Technologies. In R. A. Calvo, S. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford Handbook of Affective Computing.* Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199942237.013.032

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. C. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. http://doi.org/10.1016/j.learninstruc.2012.05.00

da Silva, L. G. Z., de Souza Gomes, L. O., & de Almeida Neris, V. P. (2020). A comparative study of users' subjective feeling collection instruments. In *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems* (pp. 1–10). Diamantina Brazil: ACM. http://doi.org/10.1145/3424953.3426642

Damasio, A. R. (2006). *Descartes' Error.* London, UK: Vintage.

Damasio, A. R. (2018). *The strange order of things: Life, feeling, and the making of the cultures.* New York: Pantheon Books.

Darwin, C. R. (1872). *The expression of the emotions in man and animals* (Original edition). Penguin Classics.

Davidson, R. J., Scherer, K. R., & Goldsmith, H. (Eds.). (2003). *Handbook of Affective Sciences.* New York, NY: Oxford University Press.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. http://doi.org/10.2307/249008

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: Three experiments. *International Journal of Human-Computer Studies*, *45*(1), 19–45. http://doi.org/10.1006/ijhc.1996.0040

Davis, M. (2009). Understanding the relationship between mood and creativity: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *108*(1), 25–38. http://doi.org/10.1016/j.obhdp.2008.04.001

De Sousa, R. (1987). *The rationality of emotion.* Cambridge, Mass: MIT Press.

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, *27*(3), 1068–1078. http://doi.org/10.1016/j.chb.2010.05.018

Derick, L., Sedrakyan, G., Munoz-Merino, P. J., Delgado Kloos, C., & Verbert, K. (2017). Evaluating emotion visualizations using AffectVis, an affect-aware dashboard for students. *Journal of Research in Innovative Teaching & Learning*, *10*(2), 107–125. http://doi.org/10.1108/JRIT-05-2017-0011

Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*, *24*(3), 766–785. http://doi.org/10.1016/j.chb.2007.04.004

Desmet, P. (2003). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In *Funology* (pp. 111–123). Springer.

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 78–89. http://doi.org/10.1016/j.jmp.2015.10.003

Dillenbourg, P. (1999). What do you mean by "Collaborative Learning"? In P. Dillenbourg (Ed.), *Collaborative learning Cognitive and computational approaches* (pp. 1–15). Amsterdam, NL: Pergamon, Elsevier Science. http://doi.org/10.1.1.167.4896

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In *Heerlen, The Netherlands: Open University of the Netherlands* (pp. 61–91).

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. *Technology-Enhanced Learning*, 3–19. http://doi.org/10.1007/978-1-4020-9827-7_1

Dillenbourg, P., Lemaignan, S., Sangin, M., Nova, N., & Molinari, G. (2016). The symmetry of partner modelling. *International Journal of Computer-Supported Collaborative Learning*, *11*(2), 227–253. http://doi.org/10.1007/s11412-016-9235-5

Dixon, T. (2003). *From passions to emotions: The creation of a secular psychological category.* Cambridge New York Melbourne: Cambridge University Press.

Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, *40*(1-2), 33–60. http://doi.org/10.1016/S0167-6393(02)00070-5

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work - CSCW '92*, 107–114. http://doi.org/10.1145/143457.143468

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., ... Sander, D. (2021). The rise of affectivism. *Nature Human Behaviour*, 1–5. http://doi.org/10.1038/s41562-021-01130-8

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*. http://doi.org/10.3389/fpsyg.2015.00621

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3/4), 169–200.

Ekman, P., & Cordaro, D. (2011). What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364–370. http://doi.org/10.1177/1754073911410740

Eligio, U. X., Ainsworth, S. E., & Crook, C. K. (2012). Emotion understanding and performance during computer-supported collaboration. *Computers in Human Behavior*, *28*(6), 2046–2054.

Ellsworth, P. C. (1994). William James and emotion: Is a century of fame worth a century of misunderstanding? *Psychological Review*, *101*(2), 222–229. http://doi.org/10.1037/0033-295X.101.2.222

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, *25*(4), 949–960. http://doi.org/10.1016/j.chb.2009.04.004

Erbas, Y., Ceulemans, E., Koval, P., & Kuppens, P. (2015). The role of valence focus and appraisal overlap in emotion differentiation. *Emotion*, *15*(3), 373–382. http://doi.org/10.1037/emo0000039

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian Inference for Psychology. *Psychonomic Bulletin & Review*, *25*(1), 5–34. http://doi.org/10.3758/s13423-017-1262-3

Ez-zaouia, M., Tabard, A., & Lavoué, E. (2020). Emodash: A dashboard supporting retrospective awareness of emotions in online learning. *International Journal of*

Human-Computer Studies, *139*, 102411. http://doi.org/10.1016/j.ijhcs.2020.102411

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, *113*(3), 464–486. http://doi.org/10.1037/0096-3445.113.3.464

Feidakis, M. (2016). A Review of Emotion-Aware Systems for e-Learning in Virtual Environments. *Formative Assessment, Learning Data Analytics and Gamification: In ICT Education*, 217–242. http://doi.org/10.1016/B978-0-12-803637-2.00011-7

Feidakis, M., Caballé, S., Daradoumis, T., Jiménez, D. G., & Conesa, J. (2014). Providing emotion awareness and affective feedback to virtualised collaborative learning scenarios. *International Journal of Continuing Engineering Education and Life-Long Learning*, *24*(2), 141. http://doi.org/10.1504/IJCEELL.2014.060154

Feidakis, M., Daradoumis, T., & Caballé, S. (2011). Endowing e-learning systems with emotion awareness. *Proceedings - 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems, INCoS 2011*, 68–75. http://doi.org/10.1109/INCoS.2011.83

Feidakis, M., Daradoumis, T., Caballé, S., Conesa, J., & Gañán, D. (2013). A Dual-Modal System that Evaluates User's Emotions in Virtual Learning Environments and Responds Affectively, 23.

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories. *Perspectives on Psychological Science*, *7*(6), 555–561. http://doi.org/10.1177/1745691612459059

Fidler, F., Thorn, F. S., Barnett, A., Kambouris, S., & Kruger, A. (2018). The Epistemic Importance of Establishing the Absence of an Effect. *Advances in Methods and Practices in Psychological Science*, *1*(2), 237–244. http://doi.org/10.1177/2515245918770407

Fiedler, K. (2004). Cycle of Theory Formation. *Personality and Social Psychology Review*, *8*(2), 123–131.

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R.*

Fischer, A. H., & Manstead, A. S. R. (2016). Social Functions of Emotion and Emotion Regulation. In I. M. Lewis, J. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (4th ed.).

Fischer, A. H., & Van Kleef, G. A. (2010). Where have all the people gone? A plea for including social interaction in emotion research. *Emotion Review*, *2*(3), 208–211. http://doi.org/10.1177/1754073910361980

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a Script Theory of Guidance in Computer-Supported Collaborative Learning. *Educational Psychologist*, *48*(1), 56–66. http://doi.org/10.1080/00461520.2012.748005

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed). Hoboken, N.J: Wiley.

Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems*, *37*(2), 239–263. http://doi.org/10.1016/S0167-9236(03)00021-6

Flake, J. K., & Fried, E. I. (2020). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *Advances in Methods and Practices in Psychological Science*, *3*(4), 456–465. http://doi.org/10.1177/2515245920952393

Fontaine, J. R. J., Gillioz, C., Soriano, C., & Scherer, K. R. (2021). Linear and non-linear relationships among the dimensions representing the cognitive structure of emotion. *Cognition and Emotion*, *0*(0), 1–22. http://doi.org/10.1080/02699931.2021.2013163

Fontaine, J. R. J., & Scherer, K. R. (2013). The global meaning structure of the emotion domain: Investigating the complementary of multiple perspectives on meaning. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp. 106–128). Oxford, UK: Oxford University Press.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–1057. http://doi.org/10.1111/j.1467-9280.2007.02024.x

Fontaine, J. R. J., Scherer, K. R., & Soriano, C. (Eds.). (2013). *Components of Emotional Meaning*. Oxford, UK: Oxford University Press.

Fox, A. S. (Ed.). (2018). *The nature of emotion: Fundamental questions* (Second edition). New York, NY: Oxford University Press.

Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The Science of Visual Data Communication: What Works: *Psychological Science in the Public Interest*. http://doi.org/10.1177/15291006211051956

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Frijda, N. H., & Scherer, K. R. (2009). Emotion definitions (psychological perspectives). In D. Sander & K. R. Scherer (Eds.), *The Oxford Companion to Emotion and the Affective Sciences*. New York, NY: Oxford University Press.

Fritz, M. A. (2015). *Reinventing the Wheel: Emotional Awareness Enhancement in Computer-Mediated Collaboration with the Dynamic Emotion Wheel* (Master Thesis). Unpublished master's thesis, Geneva University.

Fritz, M. A. (2016a). Dynamic Emotion Wheel: An Emotion Awareness Tool for Computer-Supported Collaborative Learning.

Fritz, M. A. (2016b). Real-Time Emotional Awareness in Computer-Supported Collaborative Learning: Implementing Different Graphical Representations of Self-Reported

Emotions.

Fritz, M. A., & Bétrancourt, M. (2017). Providing emotional awareness in Computer-Supported Collaborative Learning with an Emotion Awareness Tool. In *17th Biennial EARLI Conference for Research on Learning and Instruction*. Tampere, FL.

Fritz, M. A., Bétrancourt, M., Molinari, G., & Pun, T. (2015). Dynamic Emotion Wheel: An Emotional Awareness Tool for Computer-Supported Collaboration.

Fritz, M. A., & Schneider, D. K. (2019). Pensée computationnelle avec JavaScript : le cours STIC I. In *Atelier@EIAH'19. Apprentissage de la pensée informatique de la maternelle à l'Université : retours d'expériences et passage à l'échelle*. Paris, France.

Fuentes, C., Herskovic, V., RodrÃÂguez, I., Gerea, C., Marques, M., & Rossel, P. O. (2017). A systematic literature review about technologies for self-reporting emotional information. *Journal of Ambient Intelligence and Humanized Computing*, *8*(4), 593–606. http://doi.org/10.1007/s12652-016-0430-z

Garrison, D. R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education*, *13*(1), 5–9. http://doi.org/10.1016/j.iheduc.2009.10.003

Geneva Emotion Research Group. (1988). Appendix F. Labels describing affective states in five major languages. In K. R. Scherer (Ed.), *Facets of emotion: Recent research* (Version re). Hillsdale, NJ: Erlabaum.

Gentsch, K., Loderer, K., Soriano, C., Fontaine, J. R. J., Eid, M., Pekrun, R., & Scherer, K. R. (2017). Effects of achievement contexts on the meaning structure of emotion words. *Cognition and Emotion*, 1–10. http://doi.org/10.1080/02699931.2017.1287668

Gillioz, C., Fontaine, J. R. J., Soriano, C., & Scherer, K. R. (2016). Mapping emotion terms into affective space: Further evidence for a four-dimensional structure. *Swiss Journal of Psychology*, *75*(3), 141–148. http://doi.org/10.1024/1421-0185/a000180

Gilmore, R. O., Kennedy, J. L., & Adolph, K. E. (2018). Practical Solutions for Sharing Data and Materials From Psychological Research. *Advances in Methods and Practices in Psychological Science*, *1*(1), 121–130. http://doi.org/10.1177/2515245917746500

Glikson, E., Cheshin, A., & Van Kleef, G. A. (2018). The Dark Side of a Smiley: Effects of Smiling Emoticons on Virtual First Impressions. *Social Psychological and Personality Science*, *9*(5), 614–625. http://doi.org/10.1177/1948550617720269

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. *Social and Personality Psychology Compass*, *10*(10), 535–549. http://doi.org/10.1111/spc3.12267

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2022). Rstanarm: Bayesian applied regression modeling via Stan.

Graesser, A. C. (2013). Evolution of Advanced Learning Technologies in the 21st Century. *Theory Into Practice*, *52*, 93–101. http://doi.org/10.1080/00405841.2013.795446

Graesser, A. C., D'Mello, S., & Strain, A. C. (2014). Emotions in Advanced Learning Technologies. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 473–493). New York, NY: Routledge.

Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and Cognition*, *17*(2), 484–95.

Grawemeyer, B., Mavrikis, M., Holmes, W., Gutiérrez-Santos, S., Wiedmann, M., & Rummel, N. (2017). Affective learning: Improving engagement and enhancing learning with affect-aware feedback. *User Modeling and User-Adapted Interaction*, *27*(1), 119–158. http://doi.org/10.1007/s11257-017-9188-z

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences, 281–291. http://doi.org/10.1017.S0048577201393198

Gross, J. J. (Ed.). (2014). *Handbook of emotion regulation* (2. ed). New York, NY: Guilford Press.

Gross, J. J. (2015). Emotion Regulation: Current Status and Future Prospects. *Psychological Inquiry*, (26), 1–26.

Gross, T., Stary, C., & Totter, A. (2005). User-centered awareness in computer-supported cooperative work-systems: Structured embedding of findings from social sciences. *International Journal of Human-Computer Interaction*, *18*(3), 323–360.

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The Taboo Against Explicit Causal Inference in Nonexperimental Psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255. http://doi.org/10.1177/1745691620921521

Grudin, J. (1994). Computer-Supported Cooperative Work: History and Focus. *IEEE Computer*, *27*, 19–26. http://doi.org/10.1109/2.291294

Guest, O., & Martin, A. E. (2021). How Computational Modeling Can Force Theory Building in Psychological Science. *Perspectives on Psychological Science*, 174569162097058. http://doi.org/10.1177/1745691620970585

Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, *11*(3), 8–26. http://doi.org/10.1080/08923649709526970

Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work*, *11*, 411–446. http://doi.org/10.1023/A:1021271517844

Gutwin, C., Greenberg, S., & Roseman, M. (1996). Workspace awareness in real-time distributed groupware: Framework, widgets, and evaluation. *People and Computers*, 281–298. http://doi.org/10.1145/257089.257286

Haeffel, G. J. (2022). Psychology needs to get tired of winning. *Royal Society Open Science*, *9*(220099). http://doi.org/10.1098/rsos.220099

Hall, J. A., Mast, M. S., & West, T. V. (2018). *The social psychology of perceiving others accurately.*

Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition & Emotion*, *24*(1), 128–140. http://doi.org/10.1080/02699930802613828

Hareli, S., Moran-Amir, O., David, S., & Hess, U. (2013). Emotions as signals of normative conduct. *Cognition & Emotion*, *27*(8), 1395–404. http://doi.org/10.1080/02699931.2013.791615

Harley, J. M. (2015). Measuring Emotions: A Survey of Cutting-Edge Methodologies Used in Computer-Based Learning Environment Research. In *Emotions, Technology, Design, and Learning* (pp. 89–114). London, UK: Academic Press, Elsevier.

Harley, J. M., Lajoie, S. P., Frasson, C., & Hall, N. C. (2017). Developing Emotion-Aware, Advanced Learning Technologies: A Taxonomy of Approaches and Features. *International Journal of Artificial Intelligence in Education*, *27*(2), 268–297. http://doi.org/10.1007/s40593-016-0126-8

Hascher, T. (2010). Learning and Emotion: Perspectives for Theory and Research. *European Educational Research Journal*, *9*(1), 13–28. http://doi.org/10.2304/eerj.2010.9.1.13

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In G. Szwillus & J. Ziegler (Eds.), *Mensch & Computer 2003: Interaktion in Bewegung* (pp. 187–196). Wiesbaden: Vieweg+Teubner Verlag. http://doi.org/10.1007/978-3-322-80058-9_19

Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbachâ€™s Alpha for Estimating Reliability. Butâ€¦. *Communication Methods and Measures*, *14*(1), 1–24. http://doi.org/10.1080/19312458.2020.1718629

Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science*, *3*(3), 446–474. http://doi.org/10.1111/j.1756-8765.2011.01150.x

Hehman, E., & Xie, S. Y. (2021). Doing Better Data Visualization. *Advances in Methods and Practices in Psychological Science*, *4*(4), 25152459211045334. http://doi.org/10.1177/25152459211045334

Henritius, E., Löfström, E., & Hannula, M. S. (2019). University students' emotions in virtual learning: a review of empirical research in the 21st century. *British Journal of Educational Technology*, *50*(1), 80–100. http://doi.org/10.1111/bjet.12699

Hoekstra, R., Kiers, H., & Johnson, A. (2012). Are Assumptions of Well-Known Statistical Techniques Checked, and Why (Not)? *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00137

Hoemann, K., Nielson, C., Yuen, A., Gurera, J. W., Quigley, K. S., & Barrett, L. F. (2021). Expertise in emotion: A scoping review and unifying framework for individual differences in the mental representation of emotional experience. *Psychological Bulletin*, *147*(11), 1159–1183. http://doi.org/10.1037/bul0000327

Hoffmann, J. D., Brackett, M. A., Bailey, C. S., & Willner, C. J. (2020). Teaching emotion regulation in schools: Translating research into practice with the RULER approach to social and emotional learning. *Emotion*, *20*(1), 105–109. http://doi.org/10.1037/emo0000649

Hutto, D. D., Robertson, I., & Kirchhoff, M. D. (2018). A New, Better BET: Rescuing and Revising Basic Emotion Theory. *Frontiers in Psychology*, *9*. http://doi.org/10.3389/fpsyg.2018.01217

Hyniewska, S., Sato, W., Kaiser, S., & Pelachaud, C. (2019). Naturalistic Emotion Decoding From Facial Action Sets. *Frontiers in Psychology*, *9*. http://doi.org/10.3389/fpsyg.2018.02678

Immordino-Yang, M. H. (2016). *Emotions, learning, and the brain: Exploring the educational implications of affective neuroscience* (First edition). New York: W. W. Norton & Company.

Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review*, 363–370.

Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in Human–Computer interaction and usability research: Ready to deliver the promises. In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573–605). http://doi.org/10.1016/B978-044451020-4/50031-1

Jacoby, W. (2000). Loess: A nonparametric, graphical tool for depicting relationships between variables.

Jacquinot, G. (1993). Apprivoiser la distance et supprimer l'absence ? Ou les défis de la formation à distance. *Revue Française de Pédagogie*, *102*(1), 55–67. http://doi.org/10.3406/rfp.1993.1305

James, W. (1884). What is an Emotion? *Mind*, *9*(34), 188–205. Retrieved from `https://www.jstor.org/stable/2246769`

Janssen, J., & Bodemer, D. (2013). Coordinated Computer-Supported Collaborative Learning: Awareness and Awareness Tools. *Educational Psychologist*, *48*(1), 40–55. http://doi.org/10.1080/00461520.2012.749153

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior*, *27*(3), 1046–1058. http://doi.org/10.1016/j.chb.2010.06.002

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 263–280. http://doi.org/10.1007/s11412-016-9238-2

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., ... Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*(1), 125–142. http://doi.org/10.1007/s11423-014-9358-1

Järvenoja, H., Volet, S., & Järvelä, S. (2013). Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process. *Educational Psychology*, *33*(1), 31–58. http://doi.org/10.1080/01443410.2012.742334

Jézégou, A. (2010). Créer de la présence à distance en e-learning. Cadre théorique, définition, et dimensions clés. *Distances et savoirs*, *8*(2), 257–274. http://doi.org/10.3166/ds.8.257-274

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, *23*(5), 524–532. http://doi.org/10.1177/0956797611430953

Karau, S., & Williams, K. D. (1993). Interpersonal relations and group social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, *65*(4), 681–706. http://doi.org/10.1037/0022-3514.65.4.681

Keltner, D. (2019). Emotional Expression: Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*, 28.

Keltner, D., & Haidt, J. (1999). Social Functions of Emotions at Four Levels of Analysis. *Cognition & Emotion*, *13*(5), 505–521. http://doi.org/10.1080/026999399379168

Kensinger, E. A., & Ford, J. H. (2020). Retrieval of Emotional Events from Memory. *Annual Review of Psychology*, *71*(1), 251–272. http://doi.org/10.1146/annurev-psych-010419-051123

Kim, C. (2012). The role of affective and motivational factors in designing personalized learning environments. *Educational Technology Research and Development*, *60*(4), 563–584. http://doi.org/10.1007/s11423-012-9253-6

Kirkwood, A., & Price, L. (2014). Technology-enhanced learning and teaching in higher education: What is â€˜enhancedâ€™ and how do we know? A critical literature review. *Learning, Media and Technology*, *39*(1), 6–36.

http://doi.org/10.1080/17439884.2013.770404

Kirschner, P. A., Kreijns, K., Phielix, C., & Fransen, J. (2015). Awareness of cognitive and social behaviour in a CSCL environment: Self- and group awareness in CSCL. *Journal of Computer Assisted Learning*, *31*(1), 59–77. http://doi.org/10.1111/jcal.12084

Kirschner, P. A., Strijbos, J.-W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational Technology Research and Development*, *52*(3), 47–66. http://doi.org/10.1007/BF02504675

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, *5*(4), 345–379.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior*, *19*(3), 335–353. http://doi.org/10.1016/S0747-5632(02)00057-2

Kreijns, K., Kirschner, P. A., Jochems, W., & van Buuren, H. (2011). Measuring perceived social presence in distributed learning groups. *Education and Information Technologies*, *16*(4), 365–381. http://doi.org/10.1007/s10639-010-9135-7

Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social Aspects of CSCL Environments: A Research Framework. *Educational Psychologist*, *48*(4), 229–242. http://doi.org/10.1080/00461520.2012.750225

Kring, A. M., Smith, D. A., & Neale, J. M. (1994). Individual Differences in Dispositional Expressiveness: Development and Validation of the Emotional Expressivity Scale, *66*(5), 934–94.

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*(4), 596–610. http://doi.org/10.1037//0022-3514.67.4.596

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). http://doi.org/10.18637/jss.v082.i13

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. http://doi.org/10.1177/1948550617697177

Lakens, D., Adolfi, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., ... Zwaan, R. A. (2018). Justify your alpha. *Nature Human Behaviour*, *2*(3), 168–171. http://doi.org/10.1038/s41562-018-0311-x

Lakens, D., & Etz, A. J. (2017). Too True to be Bad: When Sets of Studies With Significant and Nonsignificant Findings Are Probably True. *Social Psychological and*

Personality Science, *8*(8), 875–881. http://doi.org/10.1177/1948550617693058

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. http://doi.org/10.1177/2515245918770963

Lallemand, C., & Gronier, G. (2017). *Méthodes de design UX: 30 méthodes fondamentales pour concevoir et évaluer les systèmes interactifs*.

Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The Levels of Emotional Awareness Scale: A Cognitive-Developmental Measure of Emotion. *Journal of Personality Assessment*, *55*(1-2), 124–134. http://doi.org/10.1080/00223891.1990.9674052

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *HCI and Usability for Education and Work* (Vol. 5298, pp. 63–76). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-540-89350-9_6

Laurans, G., & Desmet, P. (2012). New Directions for the Non-Verbal Measurement of Emotion In Design. In *Proceedings of 8th International Design and Emotion Conference* (p. 13). London, UK.

Lavoué, E., Kazemitabar, M., Doleck, T., Lajoie, S. P., Carrillo, R., & Molinari, G. (2020). Towards emotion awareness tools to support emotion and appraisal regulation in academic contexts. *Educational Technology Research and Development*, *68*(1), 269–292. http://doi.org/10.1007/s11423-019-09688-x

Lavoué, E., Molinari, G., & Trannoi, M. (2017). Emotional Data Collection Using Self-Reporting Tools in Distance Learning Courses. In *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)* (pp. 377–378). http://doi.org/10.1109/ICALT.2017.94

Law, E. L. C., Van Schaik, P., & Roto, V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human Computer Studies*, *72*(6), 526–541. http://doi.org/10.1016/j.ijhcs.2013.09.006

Lazarus, R. S. (1966). *Psychological stress and the coping process.* New York, NY: MacGraw-Hill.

LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences*, *114*(10), E2016–E2025. http://doi.org/10.1073/pnas.1619316114

Lee, D. M. C., Rodrigo, M. M. T., Baker, R. S. J. D., Sugay, J. O., & Coronel, A. (2011). Exploring the relationship between novice programmer confusion and achievement. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6974 LNCS*(PART 1), 175–184.

http://doi.org/10.1007/978-3-642-24600-5_21

Lee, M., Pekrun, R., Taxer, J. L., Schutz, P. A., Vogl, E., & Xie, X. (2016). Teachers' emotions and emotion management: Integrating emotion regulation theory with emotional labor research. *Social Psychology of Education*, *19*(4), 843–863. http://doi.org/10.1007/s11218-016-9359-5

Leony, D., Muñoz-Merino, P. J., Pardo, A., & Kloos, C. D. (2013). Provision of awareness of learners' emotions through visualizations in a computer interaction-based environment. *Expert Systems with Applications*, *40*(13), 5093–5100. http://doi.org/10.1016/j.eswa.2013.03.030

Levenson, R. W. (1999). The Intrapersonal Functions of Emotion. *Cognition & Emotion*, *13*(5), 481–504. http://doi.org/10.1080/026999399379159

Levenson, R. W. (2014). The Autonomic Nervous System and emotion. *Emotion Review*, *6*(2). http://doi.org/10.1016/S0031-9406(10)63010-6

Levenstein, M. C., & Lyle, J. A. (2018). Data: Sharing Is Caring. *Advances in Methods and Practices in Psychological Science*, *1*(1), 95–103. http://doi.org/10.1177/2515245918758319

Leventhal, H., & Scherer, K. R. (1987). The Relationship of Emotion to Cognition: A Functional Approach to a Semantic Controversy. *Cognition & Emotion*, *1*(1), 3–28. http://doi.org/10.1080/02699938708408361

Lewis, J. R., & Sauro, J. (2018). Item Benchmarks for the System Usability Scale. *Journal of Usability Studies*, *13*(3), 158–167.

Licklider, J., & Taylor, R. (1968). The computer as a communication device. *Science and Technology*, *76*(2).

Lieberman, M. D. (2019a). Affect labeling in the age of social media. *Nature Human Behaviour*, *3*(1), 20–21. http://doi.org/10.1038/s41562-018-0487-0

Lieberman, M. D. (2019b). Boo! The consciousness problem in emotion. *Cognition and Emotion*, *33*(1), 24–30. http://doi.org/10.1080/02699931.2018.1515726

Lieberman, M. D., Eisenberger, M. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: Affective labeling disrupts amygala activity in response to affective stimuli. *Psychological Science*, *8*(5), 421–428.

Lieberman, M. D., Inagaki, T. K., Tabibnia, G., & Crockett, M. J. (2011). Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. *Emotion*, *11*(3), 468–480. http://doi.org/10.1037/a0023503

Linnenbrink-Garcia, L., Patall, E. A., & Pekrun, R. (2016). Adaptive Motivation and Emotion in Education: Research and Principles for Instructional Design. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 228–236.

http://doi.org/10.1177/2372732216644450

Locke, E. A. (2005). Why emotional intelligence is an invalid concept. *Journal of Organizational Behavior*, *26*(4), 425–431. http://doi.org/10.1002/job.318

Lowenthal, P. R., & Snelson, C. (2017). In search of a better understanding of social presence: An investigation into how researchers define social presence. *Distance Education*, *38*(2), 141–159. http://doi.org/10.1080/01587919.2017.1324727

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., ... Halpern, B. S. (2017). Our path to better science in less time using open data science tools. *Nature Ecology & Evolution*, *1*(6), 0160. http://doi.org/10.1038/s41559-017-0160

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. http://doi.org/10.3758/s13428-016-0809-y

Lund, K., Molinari, G., Séjourné, A., & Baker, M. (2007). How do argumentation diagrams compare when student pairs use them as a means for debate or as a tool for representing debate? *International Journal of Computer-Supported Collaborative Learning*, *2*(2-3), 273–295. http://doi.org/10.1007/s11412-007-9019-z

M. Nazmul Haque Nahin, a.F., Alam, J. M., Mahmud, H., & Hasan, K. (2014). Identifying emotion by keystroke dynamics and text pattern analysis. *Behaviour & Information Technology*, *33*, 987–996. http://doi.org/10.1080/0144929X.2014.907343

MacKenzie, I. S. (2013). *Human-Computer Interaction: An Empirical Research perspective*. Waltham, MA: Morgan Kaufmann.

Makel, M. C., Hodges, J., Cook, B. G., & Plucker, J. (2019). *Questionable and Open Research Practices in Education Research* (Preprint). EdArXiv.

Makhkamova, A., Ziegler, P., & Werth, D. (2019). Augmenting Collaboration with Invisible Data: Brain-Computer Interface for Emotional Awareness. In *Proceedings of Mensch und Computer 2019 on - MuC'19* (pp. 783–787). Hamburg, Germany: ACM Press. http://doi.org/10.1145/3340764.3344908

Makowski, D., Ben-Shachar, M. S., Chen, S. H. A., & Lüdecke, D. (2019). Indices of Effect Existence and Significance in the Bayesian Framework. *Frontiers in Psychology*, *10*.

Mänty, K., Järvenoja, H., & Törmänen, T. (2020). Socio-emotional interaction in collaborative learning: Combining individual emotional experiences and group-level emotion regulation. *International Journal of Educational Research*, *102*, 101589. http://doi.org/10.1016/j.ijer.2020.101589

Marchand, G. C., & Gutierrez, A. P. (2012). The role of emotion in the learning process: Comparisons between online and face-to-face learning settings. *The Internet and Higher Education*, *15*(3), 150–160. http://doi.org/10.1016/j.iheduc.2011.10.001

Marsella, S., Gratch, J., & Petta, P. (2010a). Computational models of emotion. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for Affective Computing. A Sourcebook* (pp. 21–41). New York, NY: Oxford University Press.

Marsella, S., Gratch, J., & Petta, P. (2010b). Computational models of emotion. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for affective computing. A sourcebook* (pp. 21–41). Oxford University Press.

Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2016). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cognition and Emotion*, *30*(5), 939–952. http://doi.org/10.1080/02699931.2015.1035229

Matthews, G., Zeidner, M., & Roberts, R. D. (Eds.). (2007). *The science of emotional intelligence: Knowns and unknowns*. Oxford ; New York: Oxford University Press.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*(2), 209–237. http://doi.org/10.1080/02699930802204677

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing experiments and analyzing data: A model comparison perspective* (Third edition). New York, NY: Routledge.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The Ability Model of Emotional Intelligence: Principles and Updates. *Emotion Review*, *8*(4), 290–300. http://doi.org/10.1177/1754073916639667

Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge ; New York, NY: Cambridge University Press.

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2nd ed.). Boca Raton: Taylor and Francis, CRC Press.

McPhetres, J., Albayrak-Aydemir, N., Mendes, A. B., Chow, E. C., Gonzalez-Marquez, P., Loukras, E., … Volodko, K. (2021). A decade of theory as reflected in Psychological Science (2009–2019). *PLOS ONE*, *16*(3), e0247986. http://doi.org/10.1371/journal.pone.0247986

Meehl, P. E. (1990). Appraising and Amending Theories : The Strategy of Lakatosian Defense and Two Principles That Warrant It. *Psychological Inquiry*, *1*(2), 108–141.

Miceli, M., & Castelfranchi, C. (2019). Meta-emotions and the complexity of human emotional experience. *New Ideas in Psychology*, *55*, 42–49. http://doi.org/10.1016/j.newideapsych.2019.05.00

Michinov, N., & Michinov, E. (2008). Face-to-face contact at the midpoint of an online collaboration: Its impact on the patterns of participation, interaction, affect, and behavior over time. *Computers and Education*, *50*(4), 1540–1557. http://doi.org/10.1016/j.compedu.2007.03.002

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human*

Behavior, *52*, 573–588. http://doi.org/10.1016/j.chb.2015.01.050

Minge, M., Thüring, M., Wagner, I., & Kuhr, C. V. (2017). The meCUE Questionnaire: A Modular Tool for Measuring User Experience. In M. Soares, C. Falcão, & T. Z. Ahram (Eds.), *Advances in Ergonomics Modeling, Usability & Special Populations* (pp. 115–128). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-41685-4_11

Molinari, Gaëlle, Bozelle, C., Cereghetti, D., Chanel, G., Bétrancourt, M., & Pun, T. (2013). Feedback émotionnel et collaboration médiatisée par ordinateur: Quand la perception des interactions est liée aux traits émotionnels. In C. Choquet, P. Dessus, M. Lefevre, J. Broisin, & P. Catteau (Eds.), (pp. 305–326). Touluse, France: IRIT Press.

Molinari, Gaëlle, Chanel, G., Bétrancourt, M., Pun, T., & Bozelle, C. (2013). Emotion feedback during computer-mediated collaboration: Effects on self-reported emotions and perceived interaction. In *Computer-Supported Collaborative Learning Conference, CSCL* (Vol. 1, pp. 336–343). Madison, WI.

Molinari, GaÃÆ'NANAÃNAÂlle, Sangin, M., Dillenbourg, P., & NÃÂss, M. A. (2009). Knowledge interdependence with the partner, accuracy of mutual knowledge model and computer-supported collaborative learning. *European Journal of Psychology of Education*, *24*(2), 129–144. http://doi.org/10.1007/BF03173006

Molinari, Gaëlle, Sangin, M., Dillenbourg, P., & Nüssli, M. A. (2009). Knowledge interdependence with the partner, accuracy of mutual knowledge model and computer-supported collaborative learning. *European Journal of Psychology of Education*, *24*(2), 129–144. http://doi.org/10.1007/BF03173006

Molinari, Gaëlle, Trannois, M., Tabard, A., & Lavoué, E. (2016). EMORE-L : Un outil de reporting des émotions pour l'apprentissage à distance EMORE-L: An emotion reporting tool for distance learning. In *IHM'16: Éducation.* Fribourg, Suisse. http://doi.org/10.1145/3004107.3004126

Montagrin, A., Brosch, T., & Sander, D. (2013). Goal conduciveness as a key determinant of memory facilitation. *Emotion (Washington, D.C.)*, *13*(4), 622–8. http://doi.org/10.1037/a0033066

Moors, A. (2009). *Theories of emotion causation: A review. Cognition & Emotion* (Vol. 23). http://doi.org/10.1080/02699930802645739

Moors, A. (2014). Flavors of appraisal theories of emotion. *Emotion Review*, *6*(4), 303–307. http://doi.org/10.1177/1754073914534477

Moors, A., Ellsworth, P. C., Scherer, K. R., & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*(2), 119–124. http://doi.org/10.1177/1754073912468165

Moos, D. C., & Azevedo, R. (2009). Learning With Computer-Based Learning Environments: A Literature Review of Computer Self-Efficacy. *Review of Educational Research*, *79*(2), 576–600. http://doi.org/10.3102/0034654308326083

Mortillaro, M., & Mehu, M. (2015). Emotions: Methods of Assessment. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 519–525). Elsevier. http://doi.org/10.1016/B978-0-08-097086-8.25058-7

Mulligan, K., & Scherer, K. R. (2012). Toward a working definition of emotion. *Emotion Review*, *4*(4), 345–357. http://doi.org/10.1177/1754073912445818

Mumenthaler, C., & Sander, D. (2012). Social appraisal influences recognition of emotions. *Journal of Personality and Social Psychology*, *102*(6), 1118–35. http://doi.org/10.1037/a0026885

Murphy, N. A., & Hall, J. A. (2011). Intelligence and interpersonal sensitivity: A meta-analysis. *Intelligence*, *39*(1), 54–63. http://doi.org/10.1016/j.intell.2010.10.001

Nathanson, L., Rivers, S. E., Flynn, L. M., & Brackett, M. A. (2016). Creating Emotionally Intelligent Schools With RULER. *Emotion*, *8*(4), 305–310.

Näykki, P., Järvelä, S., Kirschner, P. A., & Järvenoja, H. (2014). Socio-emotional conflict in collaborative learning-A process-oriented case study in a higher education context. *International Journal of Educational Research*, *68*, 1–14. http://doi.org/10.1016/j.ijer.2014.07.001

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, *69*(1), 511–534. http://doi.org/10.1146/annurev-psych-122216-011836

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375–401. http://doi.org/10.1007/s11251-008-9053-x

Netzer, L., Van Kleef, G. A., & Tamir, M. (2015). Interpersonal instrumental emotion regulation. *Journal of Experimental Social Psychology*, *58*, 124–135. http://doi.org/10.1016/j.jesp.2015.01.006

Nickerson, R. S. (1999). How we knowâ€"and sometimes misjudgeâ€"what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*(6), 737–759. http://doi.org/10.1037/0033-2909.125.6.737

Nickerson, R. S. (2001). The Projective Way of Knowing: A Useful Heuristic That Sometimes Misleads. *Current Directions in Psychological Science*, *10*(5), 168–172. http://doi.org/10.1111/1467-8721.00141

Norman, D. A. (2013). *The Design of Everyday Things*. New York, NY: Basic Books.

Norman, E., & Furnes, B. (2016). The Concept of "Metaemotion": What is There to Learn From Research on Metacognition? *Emotion Review*, *8*(2), 187–193. http://doi.org/10.1177/1754073914552913

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. http://doi.org/10.1126/science.aab2374

Nuijten, M. B., Hartgerink, C. H. J., Assen, M. A. L. M. van, Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, *48*(4), 1205–1226. http://doi.org/10.3758/s13428-015-0664-2

Oatley, K., & Johnson-Laird, P. N. (2014). Cognitive approaches to emotions. *Trends in Cognitive Sciences*, *18*(3), 134–140. http://doi.org/10.1016/j.tics.2013.12.004

Ogarkova, A. (2013). Folk emotion concepts: Lexicalization of emotional experiences across languages and cultures. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning* (pp. 47–62). Oxford, UK: Oxford University Press.

Ogata, H., & Yano, Y. (2000). Combining Knowledge Awareness and Information Filtering in an Open-ended Collaborative Learning Environment. *International Journal of Artificial Intelligence in Education (IJAIED)*, *11*, 33–46.

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447. http://doi.org/10.1037/1082-989X.8.4.434

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). http://doi.org/10.1126/science.aac4716

Osgood, C. E. (1952). The Nature and measurement of meaning. *Psychological Bulletin*, *49*(3), 197–237. http://doi.org/10.1037/h0021468

Osher, D., Kidron, Y., Brackett, M. A., Dymnicki, A., Jones, S., & Weissberg, R. (2016). Advancing the Science and Practice of Social and Emotional Learning: Looking Back and Moving Forward. *Review of Research in Education*, *XX*, 1–38. http://doi.org/10.3102/0091732X16673595

Paquelin, D. (2011). La distance : Questions de proximités. *Distances Et Savoirs*, *9*(4), 565–590. http://doi.org/10.3166/ds.9.565-590

Parkinson, B. (1996). Emotions are social. *British Journal of Psychology*.

Parkinson, B. (2008). Emotions in direct and remote social interaction: Getting through the spaces between us. *Computers in Human Behavior*, *24*(4), 1510–1529. http://doi.org/10.1016/j.chb.2007.05.006

Parkinson, B. (2010). Emotions in interpersonal interactions. In K. R. Scherer, T.

Bänziger, & E. B. Roesch (Eds.), *Blueprint for Affective Computing. A Sourcebook* (pp. 131–150). New York, NY.

Parkinson, B. (2011). Interpersonal emotion transfer: Contagion and social appraisal. *Social and Personality Psychology Compass*, *5*(7), 428–439. http://doi.org/10.1111/j.1751-9004.2011.00365.x

Parkinson, B. (2013). Contextualizing Facial Activity. *Emotion Review*, *5*(1), 97–103. http://doi.org/10.1177/1754073912457230

Parkinson, B., Fischer, A. H., & Manstead, A. S. R. (2005). *Emotion in social relations: Cultural, group, and interpersonal processes.* New York: Psychology Press.

Parkinson, B., & Manstead, A. S. R. (2015). Current Emotion Research in Social Psychology: Thinking About Emotions and Other People. *Emotion Review*, *7*(4), 371–380. http://doi.org/10.1177/1754073915590624

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. http://doi.org/10.1037/0033-2909.116.2.220

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* Cambridge, U.K. ; New York: Cambridge University Press.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer.* Chichester, West Sussex: Wiley.

Pearl, J., & Mackenzie, D. (2018b). *The book of why: The new science of cause and effect.* Basic Books.

Pearl, J., & Mackenzie, D. (2018a). *The book of why: The new science of cause and effect.* Basic Books.

Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction*, *15*(5), 497–506. http://doi.org/10.1016/j.learninstruc.2005.07.014

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. http://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. http://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Positive Emotions in Education. *Beyond Coping: Meeting Goals, Visions and Challenges*, 149–173. http://doi.org/10.1207/S15326985EP3702

Pekrun, R., & Linnenbrink-Garcia, L. (Eds.). (2014a). *International Handbook of Emotions in Education.* New York, NY: Routledge.

Pekrun, R., & Linnenbrink-Garcia, L. (2014b). Introduction to Emotions in Education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 1–10). New York, NY: Routledge.

Pekrun, R., Muis, K. R., Frenzel, A. C., & Götz, T. (2018). *Emotions at school.* New York: Routledge, Taylor & Francis Group.

Pekrun, R., & Perry, R. P. (2014). Control-Value Theory of Achievement Emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education.* New York, NY: Routledge.

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2016). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition & Emotion*, 1–9. http://doi.org/10.1080/02699931.2016.1204989

Perrier, S. (2017). *Collaboration en environnement médiatisé par ordinateur : Des usages et de l'impact d'un outil de feedback émotionnel* (Master Thesis). Geneva University.

Pessoa, L. (2013). *The cognitive-emotional brain: From interactions to integration.* Cambridge, Massachusetts: The MIT Press.

Petrides, K. V., Frederickson, N., & Furnham, A. (2004). The role of trait emotional intelligence in academic performance and deviant behavior at school. *Personality and Individual Differences*, *36*(2), 277–293. http://doi.org/10.1016/S0191-8869(03)00084-9

Phielix, C., Prins, F. J., Kirschner, P. A., Erkens, G., & Jaspers, J. (2011). Group awareness of social and cognitive performance in a CSCL environment: Effects of a peer feedback and reflection tool. *Computers in Human Behavior*, *27*(3), 1087–1102. http://doi.org/10.1016/j.chb.2010.06.024

Picard, R. W. (2000). *Affective Computing.* Cambridge, MA: MIT press.

Picard, R. W. (2009). Affective computing. In D. Sander & K. R. Scherer (Eds.), *The Oxford Companion to Emotion and the Affective Sciences* (pp. 12–15). New York, NY: Oxford University Press.

Pinker, S. (1990). A theory of graph comprehension. *Artificial Intelligence and the Future of Testing*, 73–126. http://doi.org/10.1145/2046684.2046699

Plamper, J. (2015). *The history of emotions: An introduction* (First edition). Oxford, United Kingdom: Oxford University Press.

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2015). Attentional Bias for Positive Emotional Stimuli : A Meta-Analytic investigation. *Psychological Bulletin*, *142*, 79–106. http://doi.org/10.1037/bul0000026

Poole, A., & Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. *Encyclopedia of Human-*

Computer Interaction, 211–219. http://doi.org/10.4018/978-1-59140-562-7

Pucillo, F., & Cascini, G. (2014). A framework for user experience, needs and affordances. *Design Studies*, *35*, 160–179. http://doi.org/10.1016/j.destud.2013.10.001

Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J.-M. (2018). Emotion Recognition Using Physiological Signals: Laboratory vs. Wearable Sensors. In T. Ahram & C. FalcÃÂo (Eds.), *Advances in Human Factors in Wearable Technologies and Game Design* (Vol. 608, pp. 15–22). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-60639-2_2

Reeck, C., Ames, D. R., & Ochsner, K. N. (2016). The Social Regulation of Emotion: An Integrative, Cross-Disciplinary Model. *Trends in Cognitive Sciences*, *20*(1), 47–63. http://doi.org/10.1016/j.tics.2015.09.003

Reis, R. C. D., Isotani, S., Rodriguez, C. L., Lyra, K. T., Jaques, P. A., & Bittencourt, I. I. (2018). Affective states in computer-supported collaborative learning: Studying the past to drive the future. *Computers & Education*, *120*, 29–50. http://doi.org/10.1016/j.compedu.2018.01.015

Reis, R. C. D., Rodriguez, C. L., Lyra, K. T., Jaques, P. A., Bittencourt, I. I., & Isotani, S. (2015). Affective States in CSCL Environments: A Systematic Mapping of the Literature. In *2015 IEEE 15th International Conference on Advanced Learning Technologies* (pp. 335–339). http://doi.org/10.1109/ICALT.2015.95

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research.* manual, Evanston, Illinois: Northwestern University. Retrieved from `https://CRAN.R-project.org/package=psych`

Revelle, W., & Condon, D. M. (2019). Reliability from î± to ï‰: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411. http://doi.org/10.1037/pas0000754

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*(3), 410–424. http://doi.org/10.1037//0022-3514.79.3.410

Rieffe, C., Oosterveld, P., Miers, A. C., Meerum Terwogt, M., & Ly, V. (2008). Emotion awareness and internalising symptoms in children and adolescents: The Emotion Awareness Questionnaire revised. *Personality and Individual Differences*, *45*(8), 756–761. http://doi.org/10.1016/j.paid.2008.08.001

Rimé, B. (2005). *Le partage social des émotions.* Paris, France: PUF.

Rimé, B. (2009). Emotion Elicits the Social Sharing of Emotion: Theory and Empirical Review. *Emotion Review*, *1*(1), 60–85. http://doi.org/10.1177/1754073908097189

Ritchie, S. M., Hudson, P., Bellocchi, A., Henderson, S., King, D., & Tobin, K. (2016). Evolution of self-reporting methods for identifying discrete emotions in science classrooms. *Cultural Studies of Science Education*, *11*(3), 577–593.

http://doi.org/10.1007/s11422-014-9607-y

Rizzo, A. (2006). The origin and design of intentional affordances. In *Proceedings of the 6th ACM conference on Designing Interactive systems - DIS '06* (pp. 239–240). University Park, PA, USA: ACM Press. http://doi.org/10.1145/1142405.1142407

Robison, J., McQuiggan, S., & Lester, J. (2009). Evaluating the consequences of affective feedback in intelligent tutoring systems. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (pp. 1–6). Amsterdam, Netherlands: IEEE. http://doi.org/10.1109/ACII.2009.5349555

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist*, *65*(1), 1–12. http://doi.org/10.1037/a0018326

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interction Design: Beyond Human-Computer Interaction* (3rd editio). Chichester, UK: John Wiley & Sons.

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. http://doi.org/10.1177/2515245917745629

Rolls, E. T. (2014). *Emotion and decision-making explained* (First edition). Oxford ; New York, NY: Oxford University Press.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *ComputerSupported Collaborative Learning* (pp. 69–97). Berlin, Germany: Springer. http://doi.org/10.1145/130893.952914

Roseman, I. J., & Smith, C. A. (2001). Appraisal Theory. Overview, Assumptions, Varieties, Controversies. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion. Theory, Methods, Research* (pp. 3–19). New York, NY: Oxford University Press.

Rouet, J.-F., & Tricot, A. (1995). Recherche d'informations dans les systèmes hypertextes : Des représentations de la tâche à un modèle de l'activité cognitive. *Revue Sciences Et Techniques Éducatives*, *2*(3), 307–331.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Assessing Social Presence In Asynchronous Text-based Computer Conferencing, 18.

Rourke, L., & Kanuka, H. (2009). Learning in Communities of Inquiry: A Review of the Literature, 30.

Rowe, A., & Fitness, J. (2018). Understanding the Role of Negative Emotions in Adult Learning and Achievement: A Social Functional Perspective. *Behavioral Sciences*, *8*(2), 27. http://doi.org/10.3390/bs8020027

Ruiz, S., Charleer, S., Urretavizcaya, M., Klerkx, J., Fernández-Castro, I., & Duval, E.

(2016). Supporting learning by considering emotions. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16* (pp. 254–263). http://doi.org/10.1145/2883851.2883888

Russell, James A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. http://doi.org/10.1037/h0077714

Russell, James A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology*, *45*(6), 1281–1288. http://doi.org/10.1037//0022-3514.45.6.1281

Russell, James A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145. http://doi.org/10.1037/0033-295X.110.1.145

Russell, James A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion*, *23*(7), 1259–1283. http://doi.org/10.1080/02699930902809375

Sacharin, V., Schlegel, K., & Scherer, K. R. (2012). *Geneva Emotion Wheel rating study (Report).*

Salas, E., Grossman, R., Hughes, A. M., & Coultas, C. W. (2015). Measuring Team Cohesion: Observations from the Science. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 365–374. http://doi.org/10.1177/0018720815578267

Salmela, M. (2005). What is emotional authenticity? *Journal for the Theory of Social Behavior*, (35), 219–240.

Salovey, P., & Mayer, J. D. (1990). Emotional Intelligence. *Imagination, Cognition, and Personality*, *9*, 185–211. http://doi.org/10.1016/S0962-1849(05)80058-7

Sander, D. (2013). Models of Emotion. The Affective Neuroscience Approach. In J. Armony & P. Vuilleumier (Eds.), *The Cambridge handbook of human affective neuroscience* (pp. 5–53). Cambridge ; New York: Cambridge University Press.

Sander, D., Grandjean, D., & Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. *Neural Networks : The Official Journal of the International Neural Network Society*, *18*(4), 317–52.

Sander, D., Grandjean, D., & Scherer, K. R. (2018). An Appraisal-Driven Componential Approach to the Emotional Brain. *Emotion Review*, 175407391876565. http://doi.org/10.1177/1754073918765653

Sander, D., & Scherer, K. R. (Eds.). (2009). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford, UK: Oxford University Press.

Sangin, M. (2009). Peer knowledge modeling in computer supported collaborative learning.

Sangin, Mirweis, Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, *27*(3), 1059–1067. http://doi.org/10.1016/j.chb.2010.05.032

Sauro, J., & Lewis, J. R. (2011). When Designing Usability Questionnaires, Does It Hurt to Be Positive? In *Proceedings of CHI 2011* (pp. 2215–2223). Vancouver, Canada: ACM.

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research* (2nd edition). Cambridge: Morgan Kaufmann.

Scarantino, A. (2012). How to Define Emotions Scientifically. *Emotion Review*, *4*(4), 358–368. http://doi.org/10.1177/1754073912445810

Scarantino, A. (2017). How to Do Things with Emotional Expressions: The Theory of Affective Pragmatics. *Psychological Inquiry*, *28*(2-3), 165–185. http://doi.org/10.1080/1047840X.2017.1328951

Scarantino, A. (2018). The Philosophy of Emotions and Its Impact on Affective Science. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Fourth edition, Paperback edition). New York London: The Guilford Press.

Scarantino, A., & de Sousa, R. (2021). Emotion. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2021). `https://plato.stanford.edu/archives/sum2021/entries/emotion/`; Metaphysics Research Lab, Stanford University.

Scarantino, A., & Griffiths, P. (2011). Don't Give Up on Basic Emotions. *Emotion Review*, *3*(4), 444–454. http://doi.org/10.1177/1754073911410745

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020b). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses, 1–12.

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020a). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses, 1–12.

Scherer, K. R. (1982). Emotion as a process: Function, origin and regulation. *Social Science Information*, *21*(4-5), 555–570.

Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition & Emotion*, *7*(3-4), 325–355.

Scherer, K. R. (2001). Appraisal Considered as a Process of Multilevel Sequential Checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion. Theory, Methods, Research* (pp. 92–120). New York, NY: Oxford University Press.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729.

Scherer, K. R. (2007). Componential Emotion Theory Can Inform Models of Emotional

Competence. In G. Matthews, Z. Moshe, & R. D. Roberts (Eds.), *The Science of Emotiona Intelligence. Knowns and Unknowns* (pp. 101–126). New York, NY: Oxford University Press.

Scherer, K. R. (2009a). Emotion theories and concepts (psychological perspectives). In D. Sander & K. R. Scherer (Eds.), *The Oxford Companion to Emotion and the Affective Sciences*. New York, NY: Oxford University Press.

Scherer, K. R. (2009b). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, *23*(7), 1307–1351. http://doi.org/10.1080/02699930902928969

Scherer, K. R. (2010a). Emotion and emotional competence: Conceptual and theoretical issues for modelling agents. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for Affective Computing. A Sourcebook* (pp. 3–20). Oxford, UK: Oxford University Press.

Scherer, K. R. (2010b). The component process model: Architecture for a comprehensive computational model of emergent emotion. In K. R. Scherer, T. Bänziger, & E. B. Roesch (Eds.), *Blueprint for Affective Computing. A Sourcebook* (pp. 46–70). Oxford, UK: Oxford University Press.

Scherer, K. R. (2013a). Measuring the meaning of emotion words: A domain-specific componential approach. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp. 7–30). Oxford, UK: Oxford University Press.

Scherer, K. R. (2013b). The Nature and Dynamics of Relevance and Valence Appraisals: Theoretical Advances and Recent Evidence. *Emotion Review*, *5*(2), 150–162. http://doi.org/10.1177/1754073912468166

Scherer, K. R. (2019b). Studying appraisal-driven emotion processes: Taking stock and moving to the future. *Cognition and Emotion*, *33*(1), 31–40. http://doi.org/10.1080/02699931.2018.1510380

Scherer, K. R. (2019a). Studying appraisal-driven emotion processes: Taking stock and moving to the future. *Cognition and Emotion*, *33*(1), 31–40. http://doi.org/10.1080/02699931.2018.1510380

Scherer, K. R. (2019c). Towards a prediction and data driven computational process model of emotion. *IEEE Transactions on Affective Computing*, 1–1. http://doi.org/10.1109/TAFFC.2019.2905209

Scherer, K. R. (2021a). Evidence for the existence of emotion dispositions and the effects of appraisal bias. *Emotion*, *21*(6), 1224–1238. http://doi.org/10.1037/emo0000861

Scherer, K. R. (2021b). Evidence for the existence of emotion dispositions and the effects of appraisal bias. *Emotion*, *21*(6), 1224–1238. http://doi.org/10.1037/emo0000861

Scherer, K. R. (2022). Theory convergence in emotion science is timely and realistic. *Cognition and Emotion*, *36*(2), 154–170. http://doi.org/10.1080/02699931.2021.1973378

Scherer, K. R., Bänziger, T., & Roesch, E. B. (Eds.). (2010). *Blueprint for Affective Computing. A Sourcebook*. New York, NY: Oxford University Press.

Scherer, K. R., Dan, E. S., & Flykt, A. (2006). What determines a feeling's position in affective space? A case for appraisal. *Cognition and Emotion*, *20*(1), 92–113. http://doi.org/10.1080/02699930500305016

Scherer, K. R., Dieckmann, A., Unfried, M., Ellgring, H., & Mortillaro, M. (2019). Investigating appraisal-driven facial expression and inference in emotion communication. *Emotion*. http://doi.org/10.1037/emo0000693

Scherer, K. R., & Fontaine, J. R. J. (2018). The semantic structure of emotion words across languages is consistent with componential appraisal models of emotion. *Cognition and Emotion*, 1–10. http://doi.org/10.1080/02699931.2018.1481369

Scherer, K. R., & Fontaine, J. R. J. (2019). The semantic structure of emotion words across languages is consistent with componential appraisal models of emotion. *Cognition and Emotion*, *33*(4), 673–682. http://doi.org/10.1080/02699931.2018.1481369

Scherer, K. R., Fontaine, J. R. J., & Soriano, C. (2013). CoreGRID and MiniGRID: Development and validation of two short versions of the GRID instrument. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp. 523–541). Oxford, UK: Oxford University Press.

Scherer, K. R., & Grandjean, D. (2007a). Facial expressions allow inference of both emotions and their components. *Cognition & Emotion*, *22*(5), 789–801. http://doi.org/10.1080/02699930701516791

Scherer, K. R., & Grandjean, D. (2007b). Facial expressions allow inference of both emotions and their components. *Cognition & Emotion*, *22*(5), 789–801. http://doi.org/10.1080/02699930701516791

Scherer, K. R., & Meuleman, B. (2013). Human Emotion Experiences Can Be Predicted on Theoretical Grounds: Evidence from Verbal Labeling. *PLoS ONE*, *8*(3).

Scherer, K. R., & Moors, A. (2019). The Emotion Process: Event Appraisal and Component Differentiation. *Annual Review of Psychology*, *70*(1), 719–745. http://doi.org/10.1146/annurev-psych-122216-011854

Scherer, K. R., Mortillaro, M., Rotondi, I., Sergi, I., & Trznadel, S. (2018). Appraisal-driven facial actions as building blocks for emotion inference. *Journal of Personality and Social Psychology*, *114*(3), 358–379. http://doi.org/10.1037/pspa0000107

Scherer, K. R., Shuman, V., Fontaine, J. R. J., & Soriano, C. (2013). The GRID meets the Wheel: Assessing emotional feeling via self-report. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp.

281–298). Oxford, UK: Oxford University Press.

Scherer*, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, *128*, 13–35. http://doi.org/10.1016/j.compedu.2018.09.009

Schlegel, K., Fontaine, J. R. J., & Scherer, K. R. (2017). The Nomological Network of Emotion Recognition Ability. *European Journal of Psychological Assessment*, 1–12. http://doi.org/10.1027/1015-5759/a000396

Schlegel, K., & Mortillaro, M. (2018). The Geneva Emotional Competence Test (GECo): An Ability Measure of Workplace Emotional Intelligence. *Journal of Applied Psychology*, 23.

Schmidt, K. (2002). The Problem with 'Awareness'. *Computer Supported Cooperative Work*, *11*, 285–298.

Schorr, A. (2001). Appraisal. The Evolution of an Idea. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal Processes in Emotion. Theory, Methods, Research* (pp. 20–34). New York, NY: Oxford University Press.

Schwarz, N., & Clore, G. L. (2003). Mood as information: 20 years later. *Psychological Inquiry*, *14*(3-4), 296–303.

Scollon, C. N., Kim-Prieto, C., & Scollon, C. N. (2003). Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies*, *4*(1), 5–34. http://doi.org/10.1023/A:1023605205115

Seager, W. (2002). Emotional introspection. *Consciousness and Cognition*, *11*(4), 666–687. http://doi.org/10.1016/S1053-8100(02)00027-2

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(3), 294–306. http://doi.org/10.3758/CABN.4.3.294

Sherry, L. (1995). Issues in Distance Learning. *International Journal of Educational Telecommunications*, *1*(4), 337–365. Retrieved from `https://www.learntechlib.org/primary/p/8937/`

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *4*(1), 1–32. http://doi.org/10.1146/annurev.clinpsy.3.022806.0914

Shiota, M. N., Sauter, D. A., & Desmet, P. M. (2021). What are "positive" affect and emotion? *Current Opinion in Behavioral Sciences*, *39*, 142–146. http://doi.org/10.1016/j.cobeha.2021.03.007

Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., ... Yang, X. (2018). A Review of Emotion Recognition Using Physiological Signals. *Sensors*, *18*(7, 7), 2074. http://doi.org/10.3390/s18072074

Shuman, V., Clark-Polner, E., Meuleman, B., Sander, D., & Scherer, K. R. (2017). Emotion perception from a componential perspective. *Cognition and Emotion*, *31*(1), 47–56. http://doi.org/10.1080/02699931.2015.1075964

Shuman, V., Sander, D., & Scherer, K. R. (2013). Levels of valence. *Frontiers in Psychology*, *4*, 1–17. http://doi.org/10.3389/fpsyg.2013.00261

Shuman, V., & Scherer, K. R. (2014). Concepts and Structures of Emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 13–35). New York, NY: Routledge.

Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation-different emotions: How appraisals shape our emotions. *Emotion*, *7*(3), 592–600.

Sijtsma, K. (2009). On the Use, the Misuse, and the Very Limited Usefulness of Cronbachâ€™s Alpha. *Psychometrika*, *74*(1), 107–120. http://doi.org/10.1007/s11336-008-9101-0

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.2160588

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *Afex: Analysis of factorial experiments*. Manual.

Singmann, H., & Kellen, D. (2020). An Introduction to Mixed Models for Experimental Psychology. In D. H. Spieler & E. Schumacher (Eds.), *New methods in cognitive psychology* (pp. 4–31). New York, NY: Routledge.

Smith, C., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*(4), 813–838.

Smith, E. R., Seger, C. R., & Mackie, D. M. (2007). Can emotions be truly group level? Evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology*, *93*(3), 431–446. http://doi.org/10.1037/0022-3514.93.3.431

Sollaci, L. B., & Pereira, M. G. (2004). The introduction, methods, results, and discussion (IMRAD) structure: A fifty-year survey. *Journal of the Medical Library Association*, *92*(3), 364–371.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*.

Stanley, D. J., & Meyer, J. P. (2009). Two-dimensional affective space: A new approach to orienting the axes. *Emotion*, *9*(2), 214–237. http://doi.org/10.1037/a0014612

Stone, N. J., & Posey, M. (2008). Understanding coordination in computer-mediated versus face-to-face groups. *Computers in Human Behavior*, *24*(3), 827–851. http://doi.org/10.1016/j.chb.2007.02.014

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, *1*(3), 315–337. http://doi.org/10.1007/s11412-006-9660-y

TECFA. (2019). Enseigner à distance dans l'urgence. Retrieved October 26, 2021, from `https://edutechwiki.unige.ch/fr/Enseigner_%C3%A0_distance_dans_l%27urgence`

Tobii AB. (2015). Tobii Studio.

Toma, C., Yzerbyt, V., & Corneille, O. (2010). Anticipated cooperation vs. Competition moderates interpersonal projection. *Journal of Experimental Social Psychology*, *46*(2), 375–381. http://doi.org/10.1016/j.jesp.2009.11.005

Torre, J. B., & Lieberman, M. D. (2018). Putting Feelings Into Words: Affect Labeling as Implicit Emotion Regulation. *Emotion Review*, 116–124. http://doi.org/10.1177/1754073917742706

Tran, Véronique. (2004). *The influence of emotions on decision-making processes in management teams = (l'influence des émotions sur les processus de prise de décision dans les équipes de cadres)* (PhD thesis).

Tran, Veronique, Páez, D., & Sánchez, F. (2012). Emotions and Decision-Making Processes in Management Teams: A Collective Level Analysis. *Revista de Psicología Del Trabajo y de La Organizaciones*, *28*(1), 15–24.

Tu, C.-H., & McIsaac, M. (2002). The Relationship of Social Presence and Interaction in Online Classes. *American Journal of Distance Education*, *16*(3), 131–150. http://doi.org/10.1207/S15389286AJDE1603_2

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass: Addison-Wesley Pub. Co.

Tullis, T. S., & Albert, B. (2013b). *Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics* (Second edi). Waltham, MA: Morgan Kaufmann.

Tullis, T. S., & Albert, B. (2013a). *Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics* (Second edi). Waltham, MA: Morgan Kaufmann.

Turner, P. (2005). Affordance as context. *Interacting with Computers*, *17*(6), 787–800. http://doi.org/10.1016/j.intcom.2005.04.003

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and Memory. *Frontiers in Psychology*, *8*.

http://doi.org/10.3389/fpsyg.2017.01454

Vallverdù, J. (Ed.). (2015). *Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics*. IGI Global. http://doi.org/10.4018/978-1-4666-7278-9

van de Ven, N. (2017). Envy and admiration: Emotion and motivation following upward social comparison. *Cognition and Emotion*, *31*(1), 193–200. http://doi.org/10.1080/02699931.2015.1087972

Van Kleef, G. A. (2009). How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science*, *18*(3), 184–188. http://doi.org/10.1111/j.1467-8721.2009.01633.x

Van Kleef, G. A. (2010). The Emerging View of Emotion as Social Information. *Social and Personality Psychology Compass*, *4*(5), 331–343.

Van Kleef, G. A. (2017). The Social Effects of Emotions are Functionally Equivalent Across Expressive Modalities. *Psychological Inquiry*, *28*(2-3), 211–216. http://doi.org/10.1080/1047840X.2017.1338102

Van Kleef, G. A. (2018). *The interpersonal dynamics of emotion: Toward an integrative theory of emotions as social information* (First paperback edition). Cambridge, United Kingdom New York, NY Port Melbourne, VIC New Delhi Singapore: Cambridge University Press.

Van Kleef, G. A., & Fischer, A. H. (2015). Emotional collectives: How groups shape emotions and emotions shape groups. *Cognition and Emotion*, *9931*, 1–17. http://doi.org/10.1080/02699931.2015.1081349

Van Kleef, G. A., Heerdink, M. W., & Homan, A. C. (2017). Emotional Influence in Groups: The Dynamic Nexus of Affect, Cognition, and Behavior. *Current Opinion in Psychology*, *17*, 156–161. http://doi.org/10.1016/j.copsyc.2017.07.017

Van Kleef, G. A., Van Doorn, E. A., Heerdink, M. W., & Koning, L. F. (2011). Emotion is for influence. *European Review of Social Psychology*, *22*(1), 114–163. http://doi.org/10.1080/10463283.2011.627192

Vazire, S., Schiavone, S. R., & Bottesini, J. G. (2022). Credibility Beyond Replicability: Improving the Four Validities in Psychological Science. *Current Directions in Psychological Science*, *31*(2), 162–168. http://doi.org/10.1177/09637214211067779

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. http://doi.org/10.2307/30036540

Vermeeren, A., Law, E., Roto, V., Obrist, J., Marianna Hoonhout, & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: Current state and development needs. *6th Nordic Conference on Human-Computer Interaction: Extending, NordiCHI'2010*, 521–530. http://doi.org/10.1145/1868914.1868973

Vogl, E., Pekrun, R., Murayama, K., Loderer, K., & Schubert, S. (2019). Surprise, Curiosity, and Confusion Promote Knowledge Exploration: Evidence for Robust Effects of Epistemic Emotions. *Frontiers in Psychology*, *10*. http://doi.org/10.3389/fpsyg.2019.02474

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. http://doi.org/10.3758/s13423-017-1343-3

Warne, R. (2014). A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists. http://doi.org/10.7275/SM63-7H70

Wellman, H. M. (2018). Theory of mind: The state of the art. *European Journal of Developmental Psychology*, *15*(6), 728–755. http://doi.org/10.1080/17405629.2018.1435413

West, B. T., Welch, K. B., Gałecki, A. T., & Gillespie, B. W. (2015). *Linear mixed models: A practical guide using statistical software* (Second edition). Boca Raton: CRC Press, Taylor & Francis Group.

Weston, S. J., Ritchie, S. J., Rohrer, J. M., & Przybylski, A. K. (2019). Recommendations for Increasing the Transparency of Analysis of Preexisting Data Sets. *Advances in Methods and Practices in Psychological Science*, *2*(3), 214–227. http://doi.org/10.1177/2515245919848684

Wheeler, L., & Reis, H. T. (1991). Self-Recording of Everyday Life Events: Origins, Types, and Uses. *Journal of Personality*, *59*(3), 339–354. http://doi.org/10.1111/j.1467-6494.1991.tb00252.x

Williams, M. N. (2021). Levels of measurement and statistical analyses. *Meta-Psychology*, *5*. http://doi.org/10.15626/MP.2019.1916

Winne, P. H. (2015). What is the state of the art in self-, co- and socially shared regulation in CSCL? *Computers in Human Behavior*, *52*, 628–631. http://doi.org/10.1016/j.chb.2015.05.007

Wondra, J. D., & Ellsworth, P. C. (2015). An Appraisal Theory of Empathy and Other Vicarious Emotional Experiences. *Psychological Review*, *122*(3), 411–428. http://doi.org/10.1037/a0039252

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, *45*, e1. http://doi.org/10.1017/S0140525X20001685

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, *13*(5), 803–810. http://doi.org/10.1037/a0033839