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 animals 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?

You have probably guessed from the title of the thesis – the answer put forward in the present contribution is: *emotions*. In the last few decades, emotions have been widely studied from different perspectives, giving rise to the interdisciplinary field of affective sciences (e.g., Davidson, Scherer, & Goldsmith, 2003; Sander & Scherer, 2009). There is nowadays a theoretical convergence on some characteristics of emotions, which have deeply modified many of the common (mis)conceptions about what they are, why do we have them for, which causes and consequences they have, and how we can take advantage from them in various circumstances of life both at individual and collective levels (for a recent overview, see among others Adolphs & Anderson, 2018; Barrett, 2018; Damasio, 2018; Fox, 2018; Gross, 2015; Pessoa, 2013; Plamper, 2015; Rimé, 2005; Sander, 2013; Sander, Grandjean, & Scherer, 2018; Scherer, 2005; Van Kleef, 2018).

Education and learning have also been implicated in the *affective revolution*: current trends in the field of learning psychology and education sciences advocate the integration of cognitive, social and affective interactions as a means to improve both learning processes and outcomes (see M. Baker, Andriessen, & Järvelä, 2013; Pekrun & Linnenbrink-Garcia, 2014; Pekrun, Muis, Frenzel, & Götz, 2018 for recent and overarching perspectives). Be it in the classroom, in blended, or in distance learning environments, the affective components of learning are receiving greater attention at various levels of formal and informal education. There is in fact a growing consensus in considering that learning processes and outcomes influence students' affective states and, conversely, that students' affective states influence learning processes and outcomes (for a recent overview, see Brackett, Bailey, Hoffmann, & Simmons, 2019; D'Mello, 2013; Dixon, 2012; Immordino-Yang, 2016; Järvenoja, Volet, & Järvelä, 2013; Mänty, Järvenoja, & Törmänen, 2020; Pekrun, 2005; Pekrun, Vogl, Muis, & Sinatra, 2016; Petrides, Frederickson, & Furnham, 2004; R. C. D. Reis et al., 2015; Rachel Carlos Duque Reis et al., 2018).

Especially in distance learning, though, the integration of socio-emotional information is particularly challenging, since the para-verbal cues that are available in face-to-face interactions are often limited or absent (Baltes et al., 2002; Derks, Fischer, & Bos, 2008; Kreijns, Kirschner, & Vermeulen, 2013; Marchand & Gutierrez, 2012; Parkinson, 2008; Riordan & Kreuz, 2010). One of the prospected solution to this problem is the implementation of awareness tools, that is, technological artifacts implemented in the learning environment providing information about others, which is instrumental to the task at hand (Buder, 2011; Dourish & Bellotti, 1992; Janssen & Bodemer, 2013; Janssen, Erkens, & Kirschner, 2011; Kirschner, Kreijns, Phielix, & Fransen, 2015; Miller & Hadwin, 2015; K. Schmidt, 2002). Whereas, originally, awareness tools aimed preeminently at displaying whether a person was connected to the environment and, possibly, what she was doing, current trends in the field advocate for a more thorough and holistic perspective. For instance, awareness tools should resist the temptation to mimic co-located contexts, by trying to provide exactly the same information that is available in face-to-face conditions. On the contrary, awareness tools should rather focus on information that is not available – or difficult to perceive, process or remember – even face-to-face. An example of this perspective is provided by *Knowledge Awareness Tools* (Dehler, Bodemer, Buder, & Hesse, 2011; Engelmann, Dehler, Bodemer, & Buder, 2009; Sangin, Molinari, Nüssli, & Dillenbourg, 2011), that is, tools that provide information on what other members of a group know — or don't know, by conversion – about, for instance, the subject at hand. This kind of information is not readily available in face-to-face interactions, unless it is overtly manifested in the activity. Through a knowledge awareness tool, though, learners can monitor their colleagues' knowledge when needed, and take advantage from it, for instance, by asking for help, building on common ground, tailoring exchanges on the expertise of the other, and so on.

The present contribution focuses on a particular type of awareness tool, namely an *Emotion Awareness Tool* (EAT), which have received so far little, but increasing attention (Arguedas, Xhafa, Daradoumis, & Caballe, 2015; Cernea & Kerren, 2015; Chanel et al., 2016; Chen, Ma, Cerezo, & Pu, 2014; Feidakis, Caballé, Daradoumis, Jiménez, & Conesa, 2014; Fritz & Bétrancourt, 2017; Fritz, Bétrancourt, Molinari, & Pun, 2015; Lavoué et al., 2020; Molinari, Chanel, Bétrancourt, Pun, & Bozelle, 2013). An EAT combines features of an awareness tool with that of affective systems, such as affect-aware or emotion-aware technologies, through which affective information is integrated into the technology enhanced learning environment (for a recent overview see among others Arguedas, Xhafa, Daradoumis, & Caballe, 2015; R. Calvo et al., 2015; R. A. Calvo & D'Mello, 2012; Cernea & Kerren, 2015; Feidakis, 2016; Grawemeyer et al., 2017; Jason M. Harley, Lajoie, Frasson, & Hall, 2017). Such systems, though, may vary greatly in their features and objectives, for instance with respect to how they conceptualize affect and what affective dynamics they aim to enhance in the learning process.

From an affective standpoint, affective systems can be categorized according to the aims of affective computing more generally (Picard, 2000; Scherer, Bänziger, & Roesch, 2010), which Picard (2009) categorizes in four non mutually exclusive areas: (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; and (4) computational mechanisms that stimulate internal emotions or implement their regulatory and biasing functions. Furthermore, affective systems may focus on one or more phenomena within the affective spectrum, such as preferences, dispositions, moods or emotions (Scherer, 2005). Finally, they can target one or more processes within the same affective phenomenon such as elicitation, recognition, understanding, communication or regulation (Boehner, DePaula, Dourish, & Sengers, 2007; Derks, Fischer, & Bos, 2008; Eligio, Ainsworth, & Crook, 2012; Grandjean, Sander, & Scherer, 2008; Gross, 2015; Reeck, Ames, & Ochsner, 2016; Scherer, 2001, 2007; Schlegel, Grandjean, & Scherer, 2012; Siemer, Mauss, & Gross, 2007).

From the point of view of the learning process, 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, Weber, Kerren, & Ebert (2014); Craig, D'Mello, Witherspoon, & Graesser (2008); Lavoué et al. (2020); ; Lee et al. (2016); Mänty, Järvenoja, & Törmänen (2020); Näykki, Järvelä, Kirschner, & Järvenoja (2014)]. The aim of the affective system may also be inherently linked to the subject at hand, such as learning computer programming, or overtly concerned with the acquisition of affect-related skills such as social and emotional learning (Brackett, Bailey, Hoffmann, & Simmons, 2019; Osher et al., 2016), emotional competence (Scherer, 2007; Schlegel & Mortillaro, 2018), emotional intelligence (Cherniss, 2010; Hodzic, Scharfen, Ripoll, Holling, & Zenasni, 2018; Mayer, Caruso, & Salovey, 2016; Nathanson, Rivers, Flynn, & Brackett, 2016; Salovey & Mayer, 1990), or emotion self- or social-regulation (Hoffmann, Brackett, Bailey, & Willner, 2020; Järvenoja & Järvelä, 2009; Järvenoja, Volet, & Järvelä, 2013).

In this regard, the thesis proposes the implementation and assessment of a

With a provisional definition of an EAT at hand, let us go back to the case of the *hippo* and the *turtle* co-authoring a paper. Imagine that you, the *hippo*, have just typed a long paragraph in the shared document. Shortly after, through the EAT that both you and the *turtle* have at disposal, you notice that the *turtle* is feeling *angry*. What would you infer about this information? What would you infer if, in exactly the same scenario, the *turtle* felt *happy*, or *confused*, or any other state with an affective stance? Chances are that the emotional information you came to be aware of would represent a trigger for you to reorient your attention (Andriessen, Baker, & der Puil, 2011; Eligio, Ainsworth, & Crook, 2012; Posner, 2014). For instance, you can stop and think about the paragraph you have just written and ask yourself if anything in that paragraph could have elicited the affective state of your partner. Furthermore, independently that your action is the direct cause of *anger*, *happiness*, or *confusion* in your partner, the affective state of the *turtle* will probably influence the way she will behave (Van Kleef, 2010, 2018). You may expect an *angry* partner to be less receptive, motivated or open to confrontation compared to an *happy* partner (Hareli & Hess, 2010). Or you may think your *confused* partner could benefit from an explanation or a pause to reassemble her thoughts (D'Mello, Lehman, Pekrun, & Graesser, 2014). If you are able and willing to take the *turtle*'s emotional state into account, then, you can modify the course of your actions in order to act upon the affective state of your partner (Reeck, Ames, & Ochsner, 2016; Scherer, 2007). In the meantime, the very same affective state of your partner may represent a trigger for your own emotional state (Miceli & Castelfranchi, 2019; Van Kleef, 2010). For instance, the *confusion* of your partner may prompt you to feel *anger*, if you consider it to be unjustified; *surprise*, if you consider it unexpected; or *shame*, if you consider that the quality of your own contribution is undermined (Pekrun, 2006; Sander, Grandjean, & Scherer, 2018; Scherer, 2005; Schorr, 2001; Siemer, Mauss, & Gross, 2007).

The dynamics illustrated in this example depends on a complex interaction of factors, with a potentially infinite combination of cognitive, social and affective outcomes that may determine what happens next between the *hippo* and the *turtle* (Andriessen, Baker, & der Puil, 2011; M. Baker, Andriessen, & Järvelä, 2013; D'Mello & Graesser, 2012). Nonetheless, if we consider the situation as prototypical of what may happen in distance learning, we can make abstraction of specific details about the *hippo*, the *turtle* and the paper they are writing and focus on the following assumptions.

First, learning elicits emotions, that is, short-term, affectively charged experiences that have consequences on learners' behavior. For instance, Pekrun

First, emotions play an intra-personal role that influence the organism (see among others Adolphs & Anderson, 2018; Adolphs & Andler, 2018; Brosch, Scherer, Grandjean, & Sander, 2013; Frijda, 1986; Levenson, 1999; Leventhal & Scherer, 1987; Pessoa, 2013; Sander, Grandjean, & Scherer, 2018; Scherer, 1982, 2005, 2009). For instance, emotions are known to influence high-level cognitive functions such as perception, attention, memory and decision-making (Brosch, Scherer, Grandjean, & Sander, 2013), which are all implicated in learning. A functional perspective on the intra-personal role of emotions is, for instance, posited by the Component Process Model theoretical framework (Sander, Grandjean, & Scherer, 2018; Scherer, 2001, 2009) based upon the assumption that emotions have replaced reflexes to warrant a more flexible and adaptive response of the organism to the changing environment. A compatible perspective is also purported by neuroscience evidence identifying the functional role of emotions across species, independently of how emotions are *wired* into the organism (Adolphs & Anderson, 2018; Adolphs & Andler, 2018). Rather than positive or negative per se, emotions therefore represent functional or dysfunctional phenomena, that a person can either exploit to better adapt to a complex environment – as the one represented by distance, technology enhanced learning – or suffer if the emotional episode hinders the possibility to carry on the learning activity (*e.g.*, Brackett, Bailey, Hoffmann, & Simmons, 2019; D'Mello, 2013; D'Mello, Lehman, Pekrun, & Graesser, 2014; Scherer, 2007; Siemer, Mauss, & Gross, 2007).

Second, emotions play an inter-personal role in social interactions (see among others Butler, Gomes De Mesquita, & Feldman Barrett, 2017; A. H. Fischer & Manstead, 2016; A. H. Fischer & van Kleef, 2010; Keltner & Haidt, 1999; Parkinson & Manstead, 2015; Rimé, 2009; Van Kleef, 2018).

Fourth, learners involved in a computer-mediated, technology enhanced learning activity can benefit from cues about the presence of their colleagues and – especially in a collaborative setting – from a shared understanding of the situation at hand (Dillenbourg, 1999; Dillenbourg, Järvelä, & Fischer, 2009; Engelmann, Dehler, Bodemer, & Buder, 2009; Janssen & Bodemer, 2013; Janssen, Erkens, & Kirschner, 2011; Kirschner, Kreijns, Phielix, & Fransen, 2015; Kreijns, Kirschner, & Jochems, 2003; Kreijns, Kirschner, & Vermeulen, 2013; see among others Molinari, Sangin, Dillenbourg, & Nüssli, 2009; Roschelle & Teasley, 1995; Sangin, Molinari, Nüssli, & Dillenbourg, 2011; Suthers, 2006; Tu & McIsaac, 2002).

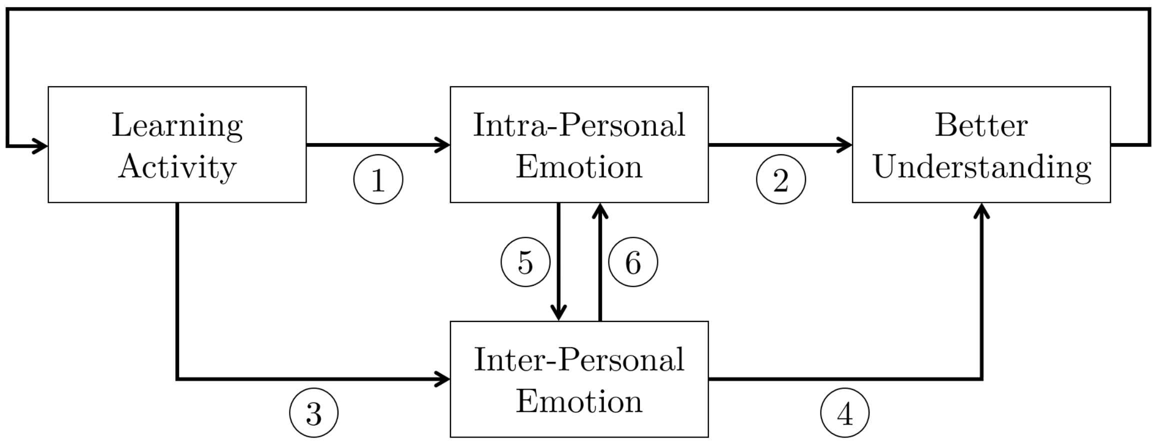


Figure 1: The EAT determines how an intra-personal emotion becomes inter-personal, and how the inter-personal emotions contribute to improve the shared understanding of the situation at hand. The situation may then trigger other intra-personal emotions.

Thesis Objectives

The present contribution therefore investigates the adoption, use and perception of a theory-driven, moment-to-moment, self-reporting emotion awareness tool in the context of distance, technology enhanced learning. The overall purpose of the thesis can be divided in two intertwined objectives: one methodological, and one empirical.

Methodological Objective

From the methodological standpoint, the way emotion awareness is expressed and perceived is a pivotal element in determining to what extent it is possible to take full advantage from the use of an EAT. In other words, the EAT is not purely an auxiliary instrument (Meehl, 1990), but an integral part of the subject matter that requires careful assessment on its own (Buder, 2011; Janssen, Erkens, & Kirschner, 2011).

In previous works, I have outlined the basis for an emotion awareness tool, the Dynamic Emotion Wheel [DEW; Fritz, Bétrancourt, Molinari, & Pun (2015); Fritz (2016a); Fritz & Bétrancourt (2017); Fritz (2015)], derived from the Geneva Emotion

Wheel [GEW; Scherer, Shuman, Fontaine, & Soriano (2013); Scherer (2005)], a widely adopted emotion self-report instrument built upon the Component Process Model theoretical framework (Sander, Grandjean, & Scherer, 2018; Scherer, 2001, 2009). The GEW (*op. cit*) comprises the graphical representation of a circumplex (Russell, 1980), in which the position of 20 subjective feelings (e.g. Joy, Relief, Shame or Anger) is determined by two orthogonal appraisal dimensions: valence, which determines to what extent the situation is pleasant or unpleasant, and control/power, which defines to what extent the situation can be acted upon by the person (Scherer, 2001; Scherer, Shuman, Fontaine, & Soriano, 2013). As a consequence, users can point to one of the 20 subjective feelings, each of them varying according to a combination of positive or negative valence, and positive or negative control/power. The GEW hence combines the two most widely adopted emotion self-report techniques – the dimensional and the discrete approaches – by implementing an emotion structure into the tool (see also Scherer, Dan, & Flykt, 2006), rather than presenting an arbitrary set of discrete natural language words or idioms with affective connotation (Scherer, 2005).

The DEW (*op. cit.*) is a web-based application that adopts the same logic, but uses the circumplex as an underlying affective space as a source from which retrieve the valence and control/power *coordinates* of each subjective feeling. At this point, the two appraisal dimensions can be used as active measures, for users to rate to what extent the situation is pleasant or unpleasant, and to what extent they can modify the current course of events. Based on the active appraisal of the affective dimensions, the DEW uses a parsimonious computational model to propose a subset of most likely subjective feelings to occur, given the rating of valence and control/power provided. The user can then decide whether one of the suggested subjective feelings corresponds to her conscious experience (Grandjean, Sander, & Scherer, 2008), in which case she can click on a button labeled with the feeling, or provide another natural language word that better represents her current emotional state.

In the thesis, I present a generalization of the DEW underlying mechanism, which I implemented into a web-based platform, the Dynamic Emotion Wheel Research Toolbox (DEW-RT). The toolbox, that can be used on- and off-line, allows researchers in affective science and related fields, as well as practitioners in education who wants to implement an affective system in their course, to set up and configure the Dynamic Emotion Wheel for their own purpose. The toolbox presents the following main features.

Self-Reported Emotions. Contrary to other affective systems that use autonomic emotion recognition, the DEW-RT is built around the principle of self-reported emotions (Jason Matthew Harley, 2015; Mortillaro & Mehu, 2015; Scherer, 2005). One the one hand, this choice implies cognitive and procedural efforts in order to express emotions (Jason Matthew Harley, 2015; Pashler, 1994), but, on the other, it also fosters a deeper implication of the person in the emotional experience. This view is consistent, for instance, with Boehenr and collaborators (2007), who advocate for an *interactional* approach to affective systems. According to this view, "the role of affective systems is not to transmit pre-existing emotional units, but to provide a resource for emotional meaning-making" (*ibid.*, p. 287).

Theory-Driven. Following Scherer's suggestion that an emotion self-report instrument should implement an emotion structure into the tool (Scherer, 2005), as the Geneva Emotion Wheel does, the DEW-RT abstacts the mechanism of sorting and sub-setting from two different structures of affective spaces: a circumplex and a Cartesian plane, which are both determined by two orthogonal affective dimensions. Even though recent evidence in the study of the structure of emotional meaning suggests that two dimensions are not enough to account for the wide variety of emotional experiences (Fontaine, Scherer, Roesch, & Ellsworth, 2007; Fontaine, Scherer, & Soriano, 2013a), in an active, self-reporting context they may still represent a trade-off between accuracy and immediacy. In the circumplex structure (Russell, 1980), the feelings are positioned alongside a circle, with an equivalent radial distance from the center. In the Cartesian plane, feelings can be placed on any combination of x and y coordinates. In both cases, researchers and practitioners can chose among different underlying affective space, determined by other affective dimensions – such as valence and arousal (Russell, 1980), or arousal and novelty (Gillioz, Fontaine, Soriano, & Scherer, 2016) – and comprising the kind and number of subjective feelings that best fit their theoretical or practical needs. Some affective spaces already exist or may be derived from the literature (e.g., Gillioz, Fontaine, Soriano, & Scherer, 2016; Pekrun, 2006; Russell, 1980; Scherer, Shuman, Fontaine, & Soriano, 2013), whereas others can be theoretically or empirically created for specific purposes (Fontaine, Scherer, & Soriano, 2013b; Scherer, Fontaine, & Soriano, 2013).

Moment-to-Moment[1]. One of the identified drawbacks of emotion self-reporting in ongoing activities is that emotions are often expressed at fixed moment during the task, which is momentarily suspended (*e.g.*, Eligio, Ainsworth, & Crook, 2012; Jason Matthew Harley, 2015). This procedure has two critical shortcomings. First, it may have disruptive consequences on the task itself, since the pause may occur in

[1] In previous works (*op. cit.*) and possibly in some parts of the present contribution, I use the term *real-time*, but I reckon that the moment-to-moment perspective is more appropriate. In fact, self-reporting implies a latency time between the moment the emotional episode occurs and the time it takes to report it into the system. Even if the DEW tries to minimize this latency time, it is inaccurate to talk about *real-time*, which rather occurs in autonomic emotion recognition

the middle of a critical situation. Second, it is not fully compatible with the temporal dimension of emotions. As emotions are limited in duration and elicited by a specific event in time, there is a loss of contextual meaning-making if emotions are self-reported with an imposed latency time. For instance, participants may report only the emotions they recall or *averaging* their emotional state over the elapsed segment. Due to its compact size and web-based nature, the DEW can be implemented alongside a wide varieties of technology enhanced activities, and even share the screen with them. As a consequence, users can express their emotions on a moment-to-moment basis (see Graesser, D'Mello, & Strain, 2014 for an overview and a list of further references), for the temporal dimension conveys meaningful and contextual information.

Graphical representations of emotions. Whereas a self-report instrument is concerned only by the expressing-displaying function, an awareness tool also implements the perceiving-monitoring function, through which the information is made available to others (Buder, 2011; K. Schmidt, 2002). In computer-mediated communication, emotions can be conveyed through a variety of media including written words, emoticons, avatars, *memes*, artworks and sounds (for an overview, see among others Berset, 2018; Blunden & Brodsky, 2020; Chen, Ma, Cerezo, & Pu, 2014; Derks, Fischer, & Bos, 2008; Glikson, Cheshin, & van Kleef, 2018; Leony, Muñoz-Merino, Pardo, & Kloos, 2013). The DEW-RT allows researchers to chose between a – limited, for the time being (Berset, 2018; Fritz, 2016b) – set of graphical representations (Hegarty, 2011), both individual and collective.

Open Science principles. The last decade has been characterized by rising concerns about replication of findings in social sciences, and psychology in particular, in what is generally referred to as the replication crisis (*e.g.*, Anvari & Lakens, 2019; Fritz, 2019; Open Science Collaboration, 2015; S. Schmidt, 2009). Solutions to increase the reproducibility and replication of scientific results, usually grouped under the umbrella term of *Open Science*, have been proposed at various stages of the research process including: the need for better substantive theories from which to draw hypothesis; clearer definition of and discrimination between concepts; the validation of instruments and experimental material; the justification of the sample size; the pre-registration of research hypothesis, protocols and planned data analyses; the full disclosure of unprocessed data and the computational steps implicated in obtaining results; more transparency in the reviewing and publication processes; and open access to scientific knowledge (see among others Chambers, 2017; Crüwell et al., 2019; Ferguson & Heene, 2012; Fiedler, 2004; Lowndes et al., 2017; Mayo, 2020; Morey et al., 2016; Nosek et al., 2015; Scheel, Tiokhin, Isager, & Lakens, 2020; Simmons, Nelson, & Simonsohn, 2012). In an effort to comply to some of these principles, the

DEW-RT is programmed in such a way that configuration of affective spaces and studies can be exported and shared with other researchers or practitioners, who can then reproduce exactly the same conditions.

Empirical Objective

The methodological intent of the DEW-RT is to provide a multipurpose EAT that can be adapted to different needs and purposes, but at the same time maintain a tight relationship with an underlying emotional structure. From an empirical standpoint, it is therefore worth investigating to what extent users can take full advantage of the main features illustrated in the previous section in an applied context. To this intent, the present contribution aims at assessing and comparing the adoption, use and perception of an EAT in two prototypical situations in distance, technology enhanced learning.

The first situation is synchronous and collaborative, and relates to the interdisciplinary field of Computer-Supported Collaborative Learning (Dillenbourg, 1999; CSCL; see Dillenbourg, Järvelä, & Fischer, 2009; Stahl, Koschmann, & Suthers, 2006; Suthers, 2006 for an overview). There is in fact a wide consensus in this field in considering that one of the main determinants of fruitful collaborative learning resides in the effort that learners put in building and updating a holistic representation of their partners, which is instrumental to a shared understanding of the task at hand Janssen & Bodemer (2013).

The second situation, by contrast, is asynchronous and individual, and relates to one of the greatest challenges in distance learning: provide learners with a social presence in the learning environment. In fact, one of the identified pitfalls in distance learning is that students are isolated and have limited contact with their peers. In this respect, technology enhanced learning represents a resource to contribute in solving the problem, but also a danger in aggravating it. On the one hand, technology can be used to alleviate the sentiment of lonliness by providing learners with perceptible cues in the environment about their colleagues. On the other hand, the increasing complexity and potential of learning technologies may also force learners in focusing on the individual task they are performing.

On the other hand, though, every assessment is done on some values or criterion of demarcation, which force to take position towards the subject at hand. Henceforth, I provide a general assumption, which drives and underlies the overall empirical objective of the thesis. I posit that an EAT serves as an affordance for learners to socially share their emotions in order to build. In more concrete words, that may translate in the following example:

Jane is prompted, by the presence of the EAT in her technology enhanced learning environment, to share her emotions with her colleague Paul because this information helps Paul to build and maintain a more accurate shared understanding of the situation at hand, which he can harness for his own activity. For the same reason, Jane is prompted to monitor and take into account Paul's emotions as instrumental information for her own activity.

Thesis Contribution

In which I list what the thesis can contribute to.

Thesis Outline

In which I set forth the organization of the manuscript.

Data and Code in Open Access

In which I provide the link to the repository containing all the data and code used to analyze the results.

Chapter 1

General Theoretical Background

The implementation and assessment of an Emotion Awareness Tool (EAT) in distance learning covers a wide spectrum of theories, pertaining often to multi- or inter-disciplinary fields. A comprehensive review of all the relevant literature is therefore beyond the scope and reach of the thesis. The general theoretical background will thus be preeminently tailored around the three elements that compose the acronyme EAT: Emotion, Awareness and Tool.

First, an EAT targets an *emotion* as the affective phenomenon of interest. It is therefore necessary to define what an emotion is, as well as mechanisms determining – at the intra- and inter-personal levels – causes and consequences of emotions, in particular with respect to learning processes. The section about emotion will mainly draw information from psychological theories of emotions, as well as contributions targeting the interplay between affective states and learning.

Second, an EAT is supposed to provide *awareness* to its users. Awareness is a multi-faceted concepts, which can be roughly mapped to two distinct – but possibly intertwined – phenomenon. On the one hand, intra-personal awareness: the mechanism by which a person becomes aware of something. On the other hand, inter-personal awareness, that is, the perception, processing and use of information about others. The section about awareness will focus mainly on the former, drawing from the literature about the importance and instrumentality of disposing of information about others in a computer-mediated environments, which requires alternatives channel of communication with respect to face-to-face conditions.

Third, an EAT is a *tool*, that is, a computational and information system implemented in a technology enhanced environment. It must therefore resolve, for instance, what kind of entities are targeted as input, when and how they are inserted into the system, as well as how and when these entities are then mapped to corresponding outputs, with various possible degrees of processing in between. The more technical section of the theoretical framework will mainyl draw from the literature about affective systems – such as affect-aware or emotion-aware systems – implemented in learning environments.

The theoretical framework will also consider some pivotal aspects of distance learning, especially in terms of learners' needs. Even though there is a growing consensus in considering that the frontiers between distance and co-located education are progressively falling apart, there is also ample evidence for the need to adapt the instructional design according to the conditions in which learning processes take place. The cognitive, social, and affective determinants of learning are situated and interact with the environment. As a consequence, there are fundamental differences whether learners' share both, either, or none of the spatial and temporal coordinates.

Finally, the theoretical framework will illustrate and discuss some key contributions in the literature, which have investigated emotion awareness with theoretical, methodological, or instrumental perspectives that are relevant to the thesis. These contributions will be referred to – by comparison and contrast – in the empirical chapters of the thesis, and it is therefore worth illustrating them at length.

1.1 Emotion

1.1.1 What Is an Emotion

1.1.2 Psychological Theories of Emotions

1.1.3 Appraisal Theories of Emotions

1.1.4 The Component Process Model Theoretical Framework

1.1.5 Intra-Personal Functions of Emotion

1.1.6 Inter-Personal Functions of Emotion

1.1.7 Emotion and Learning

1.2 Awareness

1.2.1 Intra-Personal Awareness

1.2.2 Inter-Personal Awareness

1.2.3 Awareness Tools in Computer-Mediated Environments

1.2.4 Emotional Awareness in Computer-Mediated Environments

1.3 Affective Systems

1.3.1 Affective Computing

1.3.2 Computational Models of Emotions

1.3.3 Affective Systems in Learning Environments

1.3.4 Affordance

1.4 Distance Learning

1.4.1 Synchronous vs. Asynchrounous Distance Learning

1.4.2 Individual vs. Collaborative Distance Learning

1.5 Illustration and Discussion of Key Contributions

1.5.1 Emotion Understanding and Performance During Computer-Supported Collaboration

1.5.2 Emotion Feedback During Computer-Mediated Collaboration

1.5.3 Towards Emotion Awareness Tools to Support Emotion and Appraisal Regulation in Academic Contexts

1.6 Summary

Chapter 2

The Dynamic Emotion Wheel: Emotional Structure and Competence

Being Emotional Awareness Tools (EAT) a recent subject of inquiry, there is still a lack of validated instruments to investigate the matter. This shortcoming is particularly relevant considering that the way emotional awareness is expressed and conveyed to others may be a major factor in determining and interpreting its effects. In other words, the EAT is not *simply* an auxiliary instrument (Meehl, 1990); on the contrary, it is an integral part of the subject matter and therefore requires careful assessment on its own (Buder, 2011; Janssen, Erkens, & Kirschner, 2011).

This chapter illustrates the core feature of the Dynamic Emotion Wheel (DEW), the theory-driven, self-report, and moment-to-moment emotion awareness tool that will be implemented and assessed in two distance learning settings in the empirical contributions of the thesis. The DEW was originally conceived and developed as a prototype during my internship in the EATMINT project (Fritz, 2016a, 2016b; Fritz & Bétrancourt, 2017; Fritz, Bétrancourt, Molinari, & Pun, 2015) and as the subject of my Master thesis (Fritz, 2015). A more thorough description of the conception and development can therefore be found in Fritz (2015). This chapter recalls some of the key points of the DEW, in particular its connection with the Geneva Emotion Wheel (Scherer, 2005; Scherer, Shuman, Fontaine, & Soriano, 2013) and its core feature of implementing an underlying emotional structure into the tool. Then, it extends this connection by suggesting that the use of the DEW activates the theoretical concept of emotional competence as defined by Scherer (2007), which comprises three sub-competences: (1) the appraisal, (2) the communication, and (3) the regulation sub-competences.

2.1 The Geneva Emotion Wheel

The complexity and diversity in the field of emotion theories is reflected in different approaches on how emotions should be assessed and measured (Boehner, DePaula, Dourish, & Sengers, 2007; Cernea & Kerren, 2015; Mauss & Robinson, 2009; Mortillaro & Mehu, 2015; Scherer, 2005). In a componential approach to emotion, ideally, all aspects of the synchronized and unfolding process should be taken into account (Scherer, 2005). Measuring all these aspects at once, though, requires dedicated material, some of which is difficult or impossible to deploy outside a laboratory. Self-report remains therefore a widely adopted method to ask a person about her emotional experience, either through *paper-and-pencil-like* questionnaire (e.g. Harmon-Jones, Bastian, & Harmon-Jones, 2016; Pekrun, Vogl, Muis, & Sinatra, 2016), or more elaborated self-report tools (Bradley & Lang, 1994; see among others Broekens & Brinkman, 2013; Caicedo & van Beuzekom, 2006; Desmet, 2003; Isomursu, Tähti, Väinämö, & Kuutti, 2007; Molinari, Trannois, Tabard, & Lavoué, 2016; Scherer, 2005; Siegert, Bock, Vlasenko, Philippou-Hubner, & Wendemuth, 2011). With respect to more elaborate tools in particular, Scherer (2005, p. 721) advocates some of the characteristics an emotion self-report tool should provide:

- Concentrate on the feeling component of emotion;
- Avoid the simple valence-arousal space;
- Rely on natural language labels;
- Allow the assessment of the intensity of the feeling;
- Build some emotion structure into the instrument, rather than presenting arbitrary emotion sets;
- Adopt a user-friendly graphical user interface.

Driven by these tenets, Scherer (2005) developed and refined over the years (Scherer, Shuman, Fontaine, & Soriano, 2013) one of the most widely adopted emotion self-report tool: the Geneva Emotion Wheel (GEW). For instance, the GEW is also presented in an overarching chapter (Shuman & Scherer, 2014) about emotion in the *International Handbook of Emotions in Education* (Pekrun & Linnenbrink-Garcia, 2014), which is, to the best of my knowledge, the most up-to-date and comprehensive resource about emotion in education.

The GEW is an instrumental implementation of the Component Process Model theoretical framework [CPM; Sander, Grandjean, & Scherer (2018); Scherer (2001); Scherer (2009)]. It combines the two most widely adopted approaches to emotion self-report – the dimensional and discrete emotions approaches (Scherer, 2005) – by establishing a causal link between the appraisal (Scherer, 2001) and the subjective feeling (Grandjean, Sander, & Scherer, 2008) components – that is, between the cognitive evaluation of the situation and the holistic experience that usually emerges to consciousness as a natural language word or idiom. This direct link is corroborated by recent empirical contributions investigating whether the emotional meaning of words often adopted, in different cultures, to express emotional experiences can be traced back to specific evaluations of the eliciting event (Fontaine & Scherer, 2013; Fontaine, Scherer, & Soriano, 2013b; Gentsch et al., 2017; Gillioz, Fontaine, Soriano, & Scherer, 2016; Scherer & Fontaine, 2018).

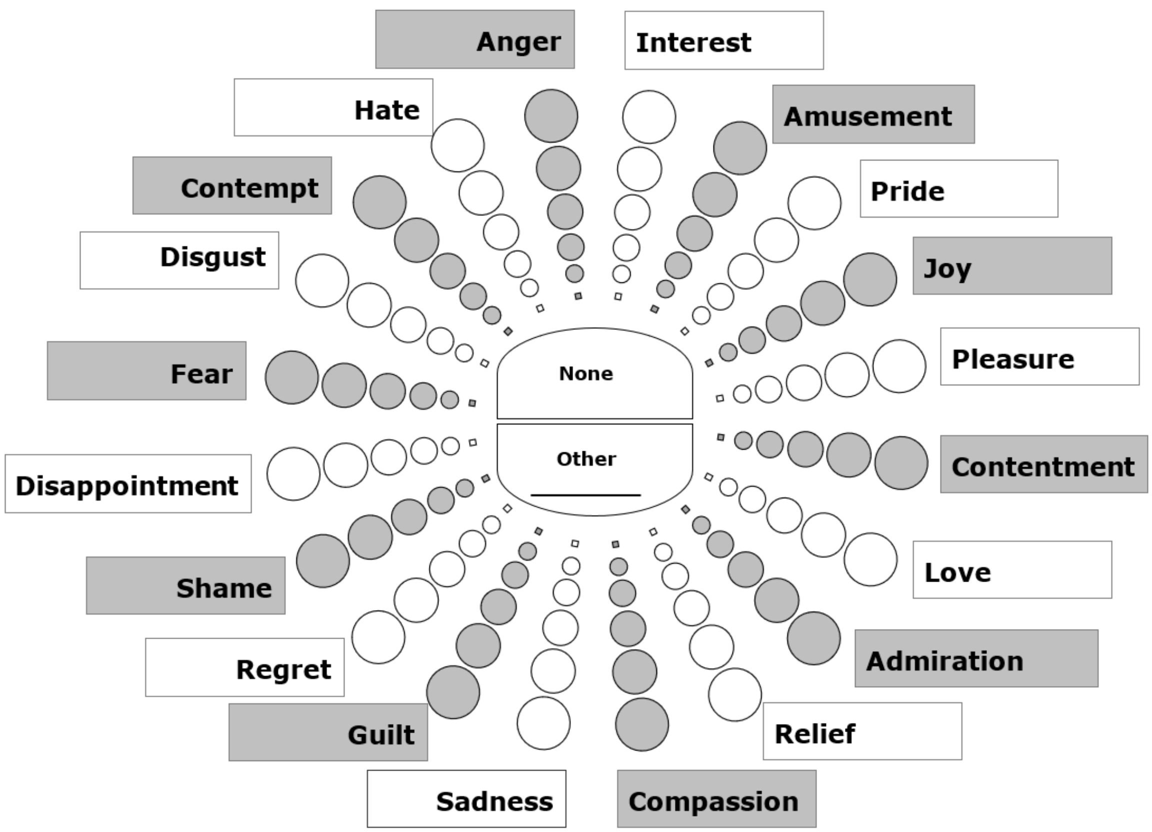


Figure 2.1: The third version of the Geneva Emotion Wheel (Scherer, Shuman, Fontaine, & Soriano, 2013)

The interface of the GEW, depicted in its third version in Figure 2.1 (Scherer, Shuman, Fontaine, & Soriano, 2013), presents 20 subjective feelings – chosen with respect to their frequency in emotion literature and everyday language use – organized in a circumplex, which is determined by two appraisal dimensions as axes: *Valence* on the abscissa and *Control/Power* on the ordinate. Each feeling is positioned according to a value along both axes and belongs thus to one of the four quadrants of the circumplex, determined by the combination of positive or negative Valence, and positive or negative Control/Power. From the origin of the axes stem 5 circles for each of the subjective feelings. 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. Users can therefore choose, first, the *row* of circles that corresponds to the subjective feeling they are experiencing and, second, the intensity of the feeling as a function of the circle's size.

2.2 The Dynamic Emotion Wheel

The GEW is a theory-driven and validated instrument that has been widely adopted in different experimental contexts. Nevertheless, from the point of view of an awareness tool, it has two major shortcomings. First, its size: the graphical interface of the GEW is too wide to be implemented alongside some other content, such as learning material or environment. Second, it is limited to self-report, and it does not provide a way for the emotional information to be stored and visualized after the expression (Berset, 2018; Hegarty, 2011; Leony, Muñoz-Merino, Pardo, & Kloos, 2013). In order to overcome these shortcomings, I took advantage of the emotional structure implemented into the GEW to build a parsimonious computational model that allows users to actively rate the two appraisal dimensions of the GEW and suggest a subset of subjective feelings that are compatible with that particular combination of appraisal dimensions.

This sections begins with a brief overview of the DEW and then depicts how the emotional structure is fit into the tool with respect to the two main functions of an awareness tool: the expressing-displaying and the perceiving-monitoring function.

2.2.1 Overview of the Dynamic Emotion Wheel

Introducing the overall interface of the DEW will ease the more detailed description of the specifics functions it serves. It is worth noting, though, that what is here presented is a configuration of the DEW that adopts and extends the same characteristics of the GEW. This is not, though, the only possible interface. The next chapter will introduce a research toolbox that precisely allows the configuration of different interfaces, which may be determined according to theoretical or practical needs of researchers interested in deploying the tool in their studies or practitioners aiming at implementing an EAT in their instructional design.

2.2.2 Emotional Structure in the Expressing-Displaying Function

2.2.3 Emotional Structure in the Perceiving-Monitroing Function

2.3 The Dynamic Emotion Wheel and Emotional Competence

Chapter 3

Introducing a Toolbox to Provide and Investigate Emotional Awareness in Computer-Mediated Environments

3.1 Configuration of Bi-Dimensional Affective Spaces

- What is an affective space
- Circumplex Affective Spaces
- Cartesian Plane Affective Spaces
- Existing Affective Spaces
- Create An Affective Space
- The EATMINT Affective Space

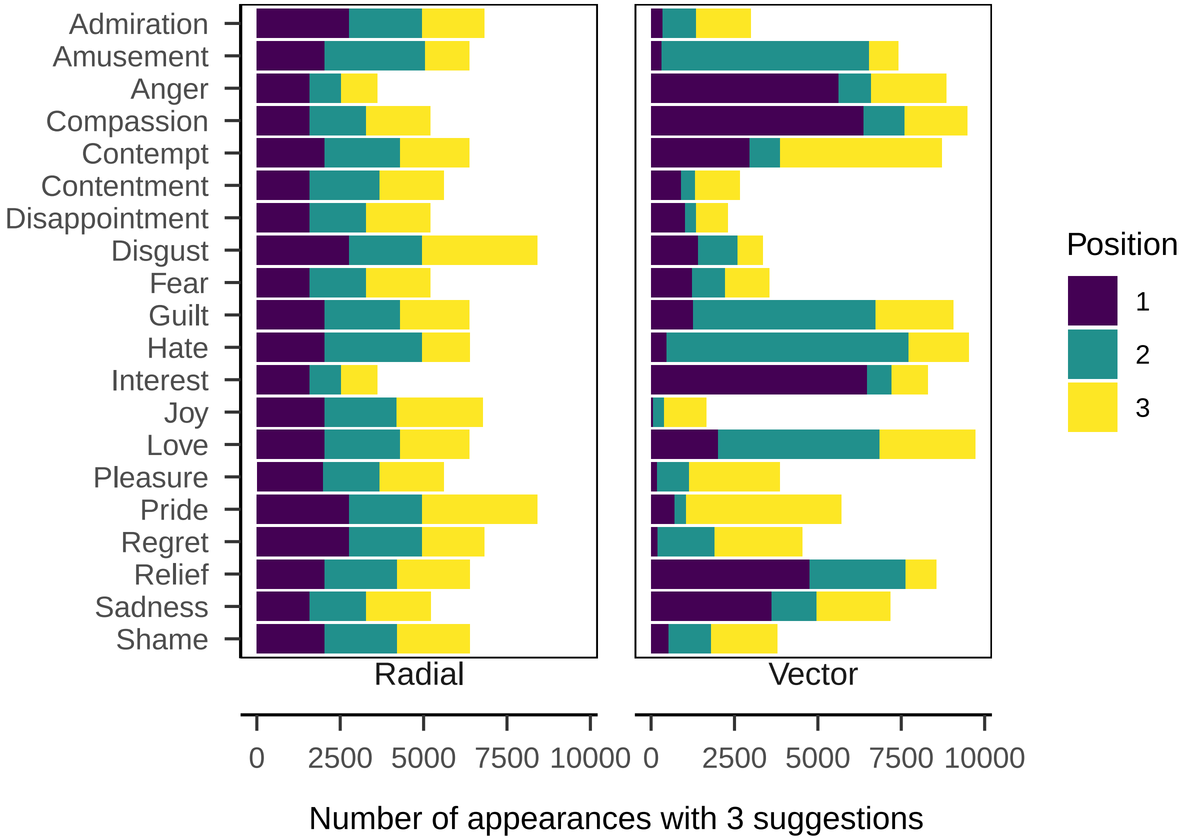
The core of the DEW is the underlying bi-dimensional affective space, which determines the relationship between the two affective dimensions and the subjective feelings. As illustrated in the previous chapter, in the case of the Geneva Emotion Wheel the two affective dimensions are Valence and Control/Power, which determine the disposition of twenty subjective feelings around a circumplex. The DEW-RT abstracts this structure in two ways. First, it provides the possibility to choose among any composition of two affective dimensions. This means that, even though the tool is deeply rooted in the Component Process Model theoretical framework (see . . .), any affective space that is based on, or can approximated to, a bi-dimensional structure can be adopted as an underlying model. Second, the DEW-RT allows any number of subjective feelings, as long as they can be mapped to the underlying bi-dimensional structure. In this regard, there are two possible structures of the affective space, which implement two different algorithms connecting the affective dimensions with the subjective feelings : (1) the circumplex structure, and (2) the Cartesian plane structure.

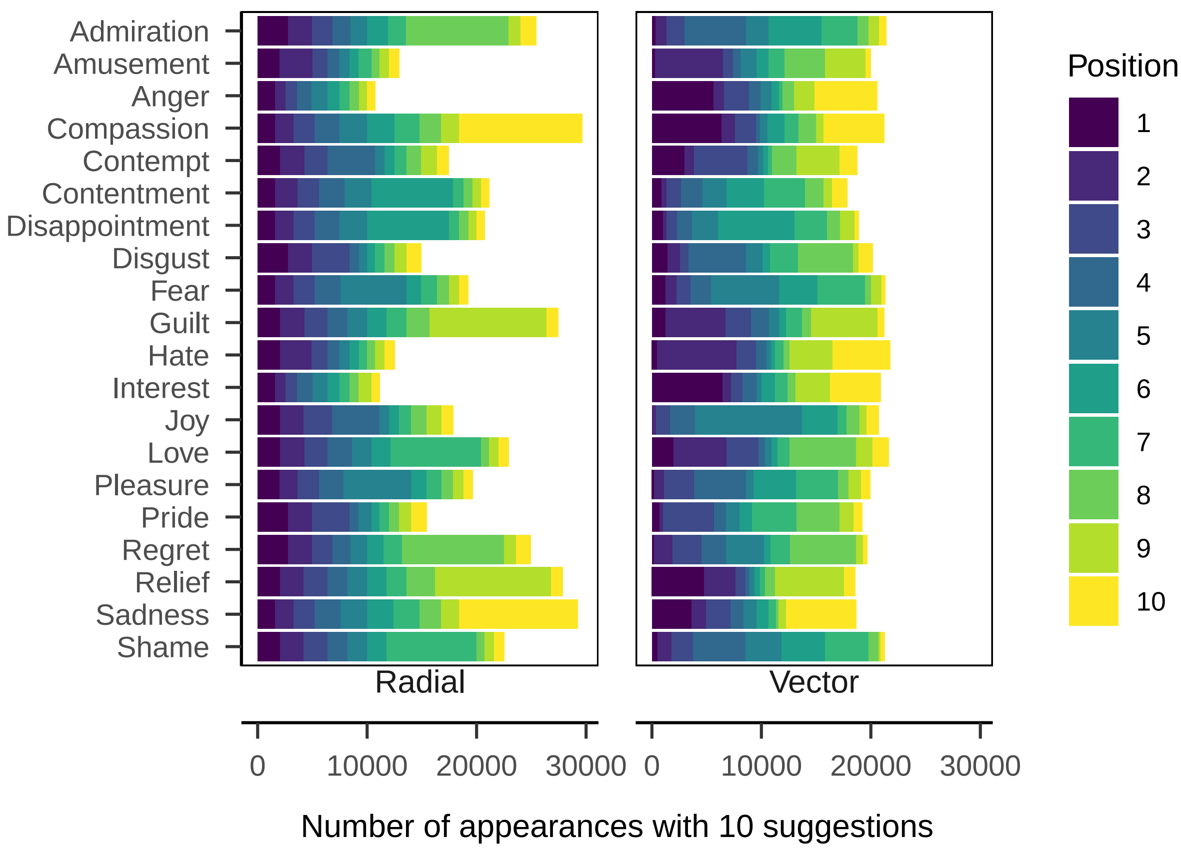
This section first proposes some examples of bi-dimensional affective spaces that can be found in the literature, and which are already available *out-of-the-box* in the DEW-RT. Next, using two existing affective spaces – one circumplex and one Cartesian plane – as sample, the section simulates the performance of the DEW subsetting algorithm using a small set of the three subjective feelings and a bigger set of ten subjective feelings. Then, the section considers how new affective spaces can be built and shared through the DEW-RT. Finally, as an example, the construction of the EATMINT circumplex, used in the two empirical contributions of the thesis, is illustrated.

3.1.1 Existing Affective Spaces

An affective space is a spatial representation of the nature of emotions. As a consequence, the affective space is defined and reflects an underlying perspective of what an emotion is.

3.1.2 Simulation of the Sub-Setting Alorithms





3.1.3 Creating and Sharing New Affective Spaces

3.1.4 The Example of the EATMINT Circumplex

3.2 Configuration of the Interface

- Displaying-Expressing Emotions (appraisal dimensions, subjective feelings)
- Monitoring-Perceiving Emotions (graphical representation of emotions)
- Exporting the Configuration

3.3 Configuration of Experimental Settings

- Login
- Task
- Data persitance
- Data exportation

Chapter 4

Emotion Awareness in Synchronous and Collaborative Settings

This chapter illustrates the first empirical contribution of the thesis, which aims at investigating the use of an Emotion Awareness Tool (EAT) in a synchronous and collaborative distance learning setting. In such a setting, learners share the temporal – but not the spatial – dimension, and also share a common learning task that must be achieved through some form of joint activity. A research field that has been investigating these learning conditions at large is Computer-Supported Collaborative Learning (CSCL) (Dillenbourg, 1999; Dillenbourg, Järvelä, & Fischer, 2009; Stahl, Koschmann, & Suthers, 2006; Suthers, 2006) – even though CSCL is not limited to distance learning, but also investigates the use of computational devices in co-located conditions. The field has, from the very beginning in the early '90, acknowledged the pivotal role of instructional designs and tools, which foster learners to build and maintain a shared understanding of the learning task (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Roschelle & Teasley, 1995; Suthers, 2006). Whereas this shared understanding was originally driven mainly from a communicative point of view (Clark & Brennan, 1991), it has later been extended to encompass a variety of sources of information, such as what a person does, thinks, wants, or even feels (Bodemer & Dehler, 2011; Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Engelmann, Dehler, Bodemer, & Buder, 2009; Kirschner, Kreijns, Phielix, & Fransen, 2015; Sangin, Molinari, Nüssli, & Dillenbourg, 2011). One of the fundamental assumptions of CSCL is, in fact, that these cues are instrumental in two processes: (1) for a learner to build and maintain a holistic representation of the colleagues with whom she shares the learning activity; and (2) for the very same colleagues to possess sufficient information to build and update a holistic representation of the learner herself. Dillenbourg and colleagues (2016) define the former 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 degrees (*e.g.* Jane thinks Paul knows that she did not understand the assignment).

It is posited, in fact, that the effort learners put to build a symmetrical representation of the partners in a collaborative task is a pivotal determinant of learning processes and, by extension, outcomes (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Molinari, Sangin, Dillenbourg, & Nüssli, 2009; Sangin, Molinari, Nüssli, & Dillenbourg, 2011). In this regard, is useful to point out two elements. First, symmetry does not mean that all the implicated learners must do, think, know, or feel the same thing. On the contrary, collaborative efforts are inherently characterized by fluctuations between socio-cognitive tensions and relaxations, which are instrumental to the learning process (Andriessen, Baker, & der Puil, 2011; Järvelä et al., 2016; Winne, 2015). Second, it is worth noting that, unlike other similar concepts implicating interpersonal accuracy (see Hall, Mast, & West, 2018 for a recent and intergative overview), mutual-modeling does not have to be persistently and precisely accurate, because that would require an effort that may take resources away from the learning activity (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 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 – a process that incidentally shares many commonalities with eliciting events in emotions (D’Mello & Graesser, 2012; Järvenoja, Volet, & Järvelä, 2013; Linnenbrink-Garcia, Patall, & Pekrun, 2016; Scherer, 2005).

CSCL literature provides broadly two non-mutually exclusive strategies to foster, among other phenomena pivotal to the effectiveness of CSCL, partner- and mutual-modeling: scripting and awareness tools (Miller & Hadwin, 2015). The first consists in accurately scaffolding the learning activity to introduce tasks through which learners can share their thoughts, doubts, experience, and so on (Dillenbourg, 2002; F. Fischer, Kollar, Stegmann, & Wecker, 2013). The second, which is more relevant to the present contribution, consists in the implementation of awareness tools, through which instrumental information about learners taking part in the joint activity is integrated into the CSCL environment, often alongside the technological artifact used to produce the outcome of the learning task (Bodemer & Dehler, 2011; Buder, 2011; Janssen & Bodemer, 2013; Kirschner, Kreijns, Phielix, & Fransen, 2015).

As mentioned in previous chapters, there is a growing consensus in considering that awareness tools should go beyond the face-to-face golden standard – that is, they should not try to provide an *ersatz* of all the verbal and para-verbal cues that are available in a co-located interaction (Buder, 2011; Kirschner, Kreijns, Phielix, & Fransen, 2015). On the contrary, awareness tools should rather provide information, particularly socially-oriented, which is not available – or is more difficult to access, process, or remember – in face-to-face settings. In this regard, even if emotions can be derived from verbal and para-verbal cues in face-to-face interaction, they require a person to be focused on her colleague's efferent manifestations (e.g. facial expressions, prosody, etc.), which is not always possible, especially when performing, in the meantime, a learning task. As a consequence, computer-mediated interaction may benefit from an alternative and dedicated way to convey emotional information (Cernea & Kerren, 2015; Derks, Fischer, & Bos, 2008; Leony, Muñoz-Merino, Pardo, & Kloos, 2013; Parkinson, 2008). As illustrated in previous chapters of the thesis, emotions provide useful information both at the intra- and inter-personal levels (Grandjean, Sander, & Scherer, 2008; Levenson, 1999; Scherer, 2005; Van Kleef, 2010, 2018)

In this empirical contribution, an EAT is therefore implemented into a computer-mediated collaborative setting as a mean to improve computer-mediated collaboration, which has already been attempted in contributions discussed at length in previous chapters. As a reminder, for instance, Eligio and colleagues (2012) compared the effect of emotional awareness – even though they a form of *scripting* rather than *awareness tool* strategy – in computer-supported collaboration for pairs of women in co-located or remote conditions, either with or without emotional awareness. They found evidence that emotional awareness improves performances in both settings. Furthermore, in the remote condition, participants also understood better the emotions of their partners, as well as experienced more positive affects compared to the co-located condition. Molinari and colleagues (2013) compared same-sex pairs collaborating through an argumentation tool in two conditions, one with an EAT and the other without, and found evidence for a beneficial effect on collaboration, even though only for pairs composed by women.

The present study builds and extends on these contributions, but with some important differences. First, the study will focus on the use of the EAT rather than on its effects on learning processes and outcomes. Second, consistently with suggested practices in the assessment of awareness tools, all participants to the study will be provided with an EAT – declined in three different versions – rather than having a control group without emotional awareness and an experimental group with emotional awareness. Finally, the collaboration will be pretended – with participants unaware of the manipulation – rather than effective. This choice serves a double purpose. On the one hand, it guarantees that all participants will be exposed to the same stimuli about the *content* of the collaboration, controlling thus sources of variance dependent on interactions between participants or the quality of the learning task. On the other hand, the use of a *unique and consistent collaborator* control for gender-related effects, which neither this study, nor the overall thesis is interested in pursuing.

More specifically, this first empirical contribution investigates the reasons why a learner may be interested in using an Emotion Awareness Tool (EAT) with respect both to the emotional information it prompts to share and the information it provides with. In this intent, three different interfaces of an EAT will be compared, which differ in how socially-oriented they are by varying (1) whether the emotional information will be disclosed to the partner or remain visible only to the learner herself; and (2) whether the learner has access to her own emotional information, that of the partner, or both. Using a randomized trial and controlled environment, in which participants – unaware of it – will collaborate with a simulated partner in a joint problem-solving tasks, performance-based indicators of the use fo the EAT will be collected: number of emotions expressed for the interest in sharing emotions, and eye-tracking measures for the interest in seeking and processing emotional information. In comparing the three interfaces, it is posited that the more socially-oriented the interface is, the greater use of the EAT in expressing, seeking and processing the emotional information will be done.

The study also provides corollary analyses linked to *dynamic* and *moment-to-moment* use of the EAT during a computer-mediated task, such as the transitions of users' gaze between different zones of the interface and the evolution of the emotional episodes expressed by participants over time. Finally, taking advantage of the fact that the experiment uses the same experimental task as in Fritz (2015), an internal meta-analysis is also provided to assess reliability of measures as reference for future implementations.

4.1 Study Overview

The moment-to-moment use of a voluntary self-report EAT implies that both the expressing-displaying and the perceiving-monitoring functions of an awareness tool (Buder, 2011; K. Schmidt, 2002) require an effort that is not directly implemented in the content space, but is limited to the relational space of the task (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Janssen & Bodemer, 2013). As pointed out in the general theoretical framework of the thesis, expressing and perceiving emotions require cognitive effort, especially when (1) emotions are expressed deliberately and at a conceptual level, as stated by the Component Process Model theoretical framework (Sander, Grandjean, & Scherer, 2018; Scherer, 2009), and (2) perceived emotions are linked to inferential processes about causes and consequences of behavior, as stated by the Emotion As Social Information (EASI) model (Van Kleef, 2009, 2018). One of the pivotal aspects of voluntarily, self-reported emotional awareness is thus whether learners are keen to make this effort and under which conditions they are more prone to dedicate some of their cognitive resources for an activity outside the content space of the task (Janssen & Bodemer, 2013; Pashler, 1994).

More broadly, this phenomenon relates to one of the fundamental questions in theory of emotion: *why do people express and share emotions* (e.g., Darwin, 1872; A. H. Fischer & Manstead, 2016; Van Kleef, 2018)? Two main positions on the matter are relevant to the context of emotional awareness provided during a computer-mediated collaborative task.

On the one hand, there is an intra-personal perspective which is, for instance, represented by the concept of *affect labeling* (Matthew D. Lieberman et al., 2007; Matthew D. Lieberman, 2019; Torre & Lieberman, 2018), according to which putting feelings into words is a form of emotion self-regulation (Gross, 2002, 2015), which may even be implicit, since the person may not realize the regulation is taking place. Similarly, one of the main tenet of the RULER approach (Brackett, Bailey, Hoffmann, & Simmons, 2019; Nathanson, Rivers, Flynn, & Brackett, 2016) – and emotional intelligence (Mayer, Caruso, & Salovey, 2016; Salovey & Mayer, 1990) or competence (Scherer, 2007) more broadly – resides in the individual ability to appraise the situation and recognize her own emotional episodes to better cope with the situation. Expressing emotions would therefore help learners to ask themselves how they are feeling by evaluating the situation (Erbas, Ceulemans, Koval, & Kuppens, 2015; Hoffmann, Brackett, Bailey, & Willner, 2020; Lavoué et al., 2020) and finding the corresponding subjective feeling that best depict their emotional state (Grandjean, Sander, & Scherer, 2008; Scherer, 2009). Applied to the computer-mediated collaborative context, learners could therefore use an EAT because it provides an affordance to differentiate and reflect about how they are feeling in a self-centered process (Boehner, DePaula, Dourish, & Sengers, 2007), aimed at regulating the emotional episodes elicited by the socio-cognitive conflicts of the learning task (Andriessen, Baker, & der Puil, 2011; Arguedas, Daradoumis, & Xhafa, 2016). As a result, learners can benefit from emotional self-awareness, even in the absence of communication with the partner.

On the other hand – even if the two positions are not mutually exclusive – many researchers advocate that emotions have pivotal inter-personal functions (Parkinson, 1996; Parkinson & Manstead, 2015; Rimé, 2009; Van Kleef, 2009, 2018; Van Kleef, Van Doorn, Heerdink, & Koning, 2011). For instance, Rimé (2009) posits that an individualistic view of emotion and regulation is untenable, and that the usefulness of emotions resides in the social sharing of them. According to this view, a person shares her emotions to make others aware of her emotional state as a way to gain social attention, arouse empathy, stimulate bonding or strengthen social ties (ibid.). Parkinson (2011) introduces the concept of *social appraisal*, according to which emotions in others can serve as useful information to modify our own behavior. An integrated framework of the inter-personal function of emotions is posited by Van Kleef in the Emotion As Social Information (EASI) model (Van Kleef, 2009, 2018), according to which emotional expression can modify behavior in others through affective reactions or inferential processes about elicited emotions' causes and consequences. For instance, the emotion of a person can represent a trigger for an emotional episode in another person, a phenomenon also known as meta-emotion (Miceli & Castelfranchi, 2019). Applied to the computer-mediated collaborative context, thus, learners can therefore use the EAT to sustain the mutual-modeling activity to foster collaborative learning (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016), especially in the absence of para-verbal cues that are usually available in face-to-face settings (Derks, Fischer, & Bos, 2008). Learners can benefit from emotional awareness by adapting their behavior according to the crossed-over emotional information available, for instance by engaging in inter-personal emotion regulation (Netzer, van Kleef, & Tamir, 2015; Reeck, Ames, & Ochsner, 2016; Zaki & Williams, 2013).

In the few studies that have investigated emotional awareness in computer-mediated collaboration so far (*e.g.*, Avry, Molinari, Bétrancourt, & Chanel, 2020; Eligio, Ainsworth, & Crook, 2012; Molinari, Chanel, Bétrancourt, Pun, & Bozelle, 2013), participants who disposed of emotional awareness (*e.g.*, those in the *treatment* group) had access both to their own emotions and that of the partner, knowing beforehand that the emotional information would be crossed-over. This ecological situation is consistent with the mutual-modeling perspective (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016), 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. It is nevertheless possible to break down this *full* use of an EAT by varying (1) the sender of emotional information provided, and (2) the recipient of the emotional information expressed (Van Kleef, 2018). As suggested by Buder (2011), varying the characteristics of the same awareness tool – rather than comparing a *control* group without the awareness tool and a *treatment* group with the tool – allows a better assessment of its contribution. For instance, by varying the sender and the recipient of the emotional information, three different interfaces of the tool can be obtained:

- *Self-Centered*: the learner can access only her own emotions, and the emotions she expresses are not conveyed to the partner;
- *Partner-Oriented*: the learner can access only the partner's emotions and the emotions she expresses are conveyed to the partner, even though she can't access them herself;
- *Mutual-Modeling*: the learner can access both her own and the partner's emotions, which provide direct and persistent comparison fostering symmetry in the emotional information shared by both partners.

The three versions of the EAT imply different reasons why a learner may be keen to use the EAT to express her own emotions, as well as to seek and process the emotional information provided by the tool. In fact, the interfaces mobilize different theoretical concepts associated with an intra- or inter-personal use of emotional awareness (see Figure 4.1 for a graphic representation). The *Self-Centered*, linked to the intra-personal function of emotions, activates concepts such as affect labeling, appraisal and self-regulation; the *Partner-Oriented*, related more towards an inter-personal perspective, implies meta-emotions, social-regulation, and social-sharing; and, finally, the *Mutual-Modeling* combines both perspectives by adding symmetry in the emotional information available to both partners. Comparing the three versions of the EAT will thus contribute to determine whether the more *socially-oriented* interface yields a more thorough use of emotional awareness through the EAT.

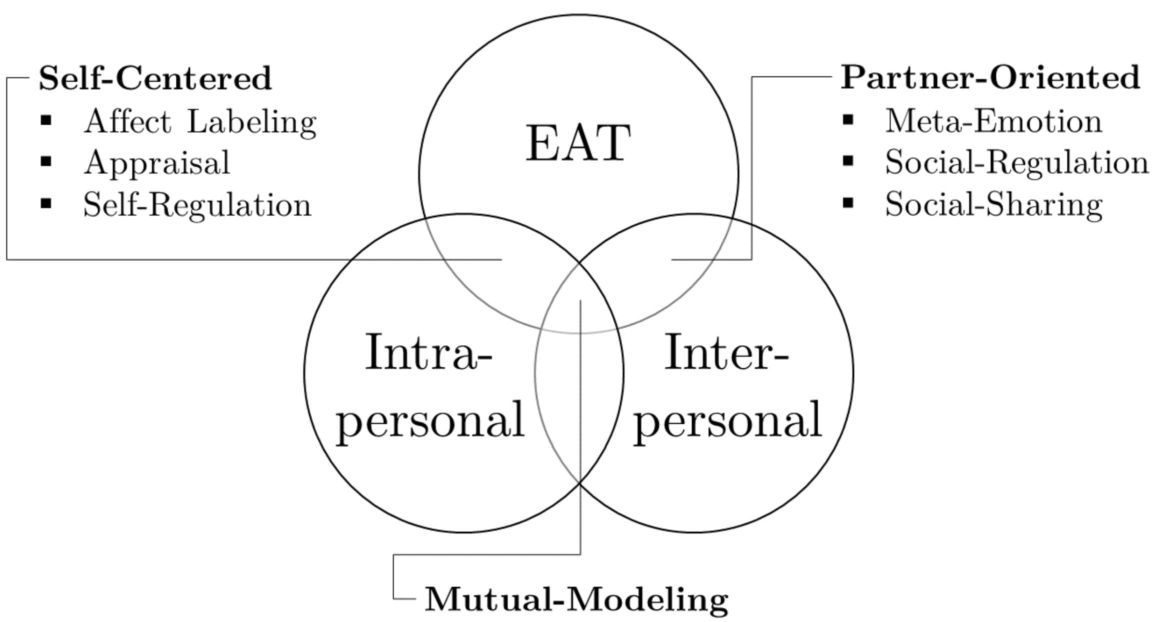


Figure 4.1: Theoretical concepts mobilized by versions of the EAT differing in the use of, and access to emotional information.

4.2 Research Question and Hypothesis

The phenomenon under scrutiny in the present study is therefore whether a different use of, and access to emotional information elicit a different use of the EAT in terms of its expressing-displaying and the perceiving-monitoring functions (Buder, 2011; K. Schmidt, 2002). More specific hypotheses are stated with respect to each function.

4.2.1 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 below the hypothesis about the use in perceiving emotions). 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 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* interface (*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. Translated in concrete use:

- Jane will express more emotions when she knows she is sharing her emotions with Paul compared to when she is expressing emotions only for herself;
- Jane will express even more emotions when she knows she is sharing her emotions with Paul *and* they can mutually dispose of direct and persistent comparison between their respective emotional episodes.

4.2.2 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. Translated in concrete use:

- Jane will seek more often and process the emotional information longer when she can access Paul's rather then her own emotions;
- Jane will seek even more often and process the emotional information even longer when she can access, for direct and persistent comparison, both her own and Paul's emotions at the same time.

4.3 Methods

I report how I determined the sample size, all data exclusions (if any), all manipulations, and all measures in the study.

4.3.1 Participants and Design

48 participants (29 women, 19 men), aged 18 to 55 ($M = 37.3, SD = 10.01$), voluntarily participated to the study. The sample size was determined by time constraints, since data had to be collected in 15 days. 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.

4.3.2 Material

Overall Interface of the CSCL Task

The experimental material comprises different components; I therefore provide an overview before specifying the various details. Figure 4.2 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.

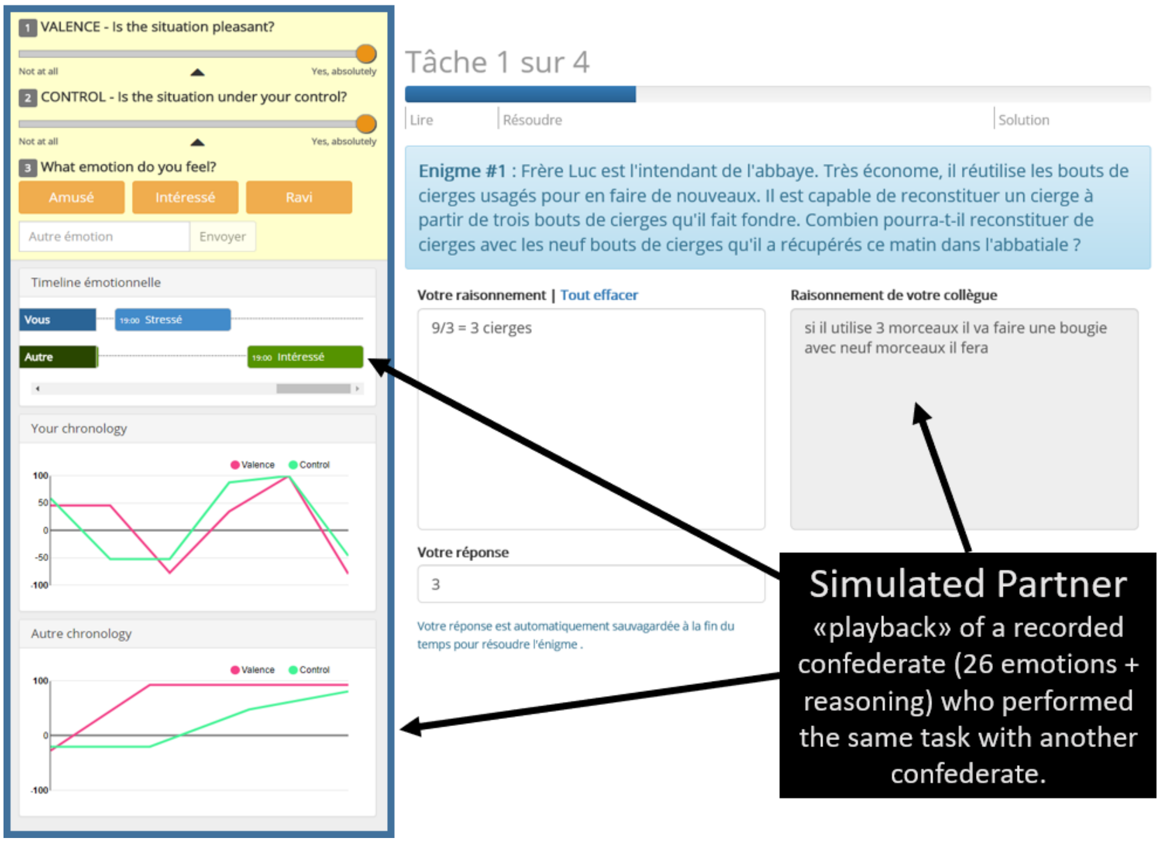


Figure 4.2: Overview of the interface that presents the various part of the material adopted for the CSCL task.

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 a previous study (Fritz, 2015), where they elicited different emotions both in number and kind in a population similar to that of the current sample. Each enigma was subdivided in three phases:

- 40 seconds during which the text of 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 answer) of the simulated partner;
- 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 version of the DEW was composed as follows. The Valence and Control/Power dimensions (Scherer, 2005; Scherer, Shuman, Fontaine, & Soriano, 2013) determined the appraisals for the expression-displaying of emotional episodes. Valence was prompted with the question "*Is the situation pleasant?*" whereas Control/Power 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*. The underlying affective space was configured using a radial circumplex with the 20 EAT-MINT emotions also used in Fritz (2015). The composition of the circumplex has been detailed in the previous chapter (add reference here).

For the monitoring/perceiving function of the EAT, each condition differed in the following ways (depicted in Figure 4.3):

- *Self-Centered*: the interface comprises an emotion timeline, then a graphic 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 is the same as in the *Self-Centered* condition, but the information provided is based only on the emotions expressed by the simulated partner (see below).
- *Mutual-Modeling*: the interface comprises an emotion timeline, 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 presents a tag cloud 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.

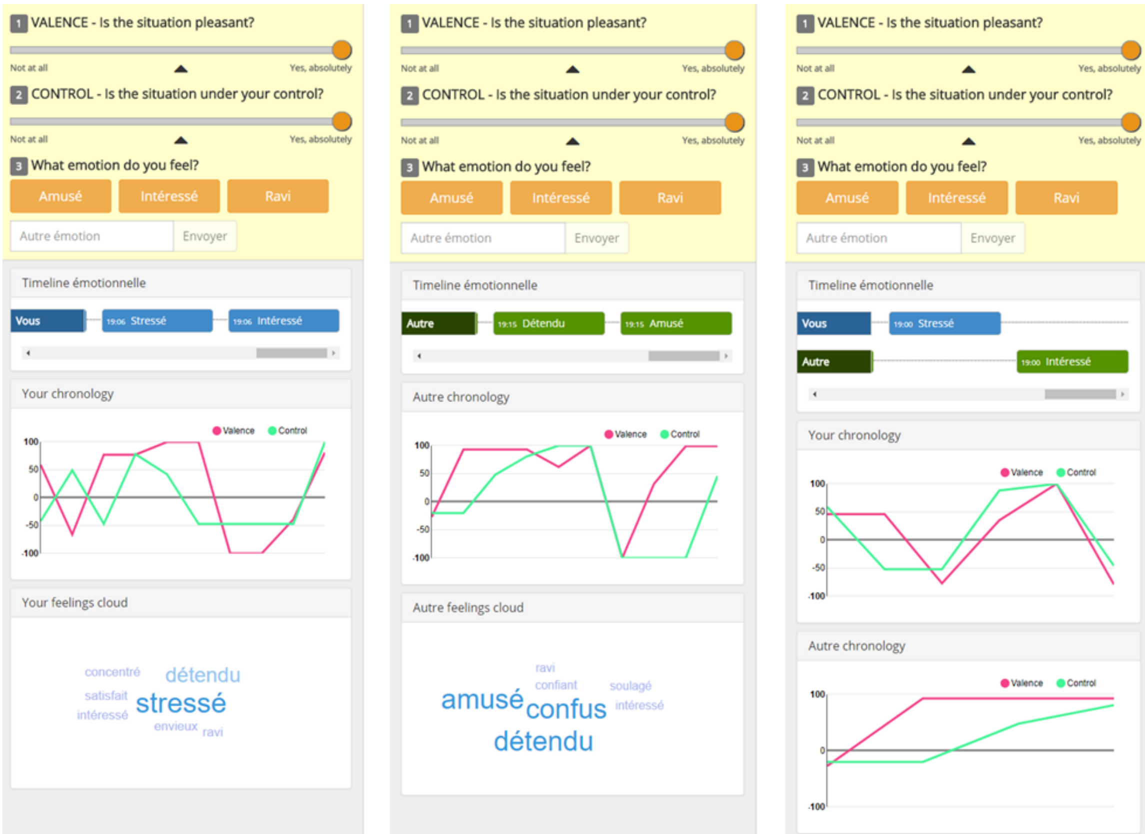


Figure 4.3: The three different interfaces used in the experiment. In order from left to right: the *Self-Centered*, the *Partner-Oriented*, and the *Mutual-Modeling* versions.

Simulated partner

The *playback* manipulations displayed on the interface were recorded in a pilot test with 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, with both the evaluation on the appraisal dimensions and the related subjective feeling expressed; and (2) what confederates 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 4.1.

Table 4.1: 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 displaying/expressing upper zone, and the monitoring/perceiving lower zone.

4.3.3 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 and introduced him 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 showed 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 study 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 which emotional information they would have access (*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.

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 a sort of *log-in form*, providing a random ID and an identifier for the pair. Once the task properly 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 with information that will not be used in the present contribution, but have been analyzed by 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.

4.3.4 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.

4.3.5 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.

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 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). Second, 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. 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. For both measures, a omnibus one-way ANOVA with pairwise comparison between all conditions were planned beforehand. A family-wise correction to account for inflation in Type I error has been planned for each pairwise comparison.

Within this experimental setting, a sensitivity analysis reveals that, given the planned sample $N = 48$ and conventional $\alpha = 0.05$ and $\beta = 0.8$ error rates control, each statistical test will be able to detect an effect size of Cohen's $d = 0.93$. Taking a previous study using exactly the same task as benchmark, that would translate in practical terms as follows:

- For expressing emotions, the statistical test will detect a difference of 6.58 expressed emotions or greater, given a mean of $M = 18.81$ ($SD = 7.10$) observed in Fritz (2015)

- For information processing, the statistical test will detect a difference of 19.92 seconds or greater, given a mean of $M = 41.11$ ($SD = 21.50$) observed in Fritz (2015);
- For information seeking, the statistical test will detect a difference of 32.08 number of visits or greater, given a mean of $M = 72.29$ ($SD = 34.62$) observed in Fritz (2015).

All analysis were conducted using the statistical software R, version 4.0.0. Data and code are available in the supplementary material.

4.4 Results

4.4.1 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 4.2.

Table 4.2: 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 the planned test and make it more exposed to violation of assumptions (Hoekstra, Kiers, & Johnson, 2012), which are checked in Appendix A. The interpretation of results must therefore take into account these limits.

4.4.2 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). 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 4.4).

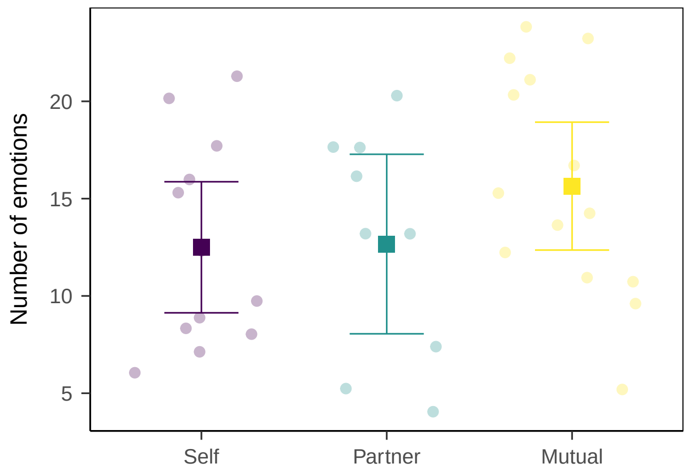


Figure 4.4: Number of emotions expressed by experimental condition. Bars represent 95% confidence intervals.

No detectable difference, neither in the omnibus one-way ANOVA ($F(2, 32)$ = 1.25, MSE = 31.82, p = .301, $\hat{\eta}^2_G$ = .072), nor in the pairwise comparisons could be observed. Results of the comparisons are depicted in Table 4.3. Hypothesis (*H1*) is therefore rejected: a different access to and use of emotional information did not yield detectable differences in the number of emotion expressed.

Table 4.3: Pairwise comparisons of the thee conditions with respect to the number of emotions expressed (p-values are adjusted with the Tukey method).

contrast	estimate	SE	df	t.ratio	p.value	cohens.d
Self - Partner	-0.17	2.49	32	-0.07	.998	-0.03
Self - Mutual	-3.14	2.22	32	-1.42	.345	-0.56
Partner - Mutual	-2.98	2.41	32	-1.23	.442	-0.53

4.4.3 Differences in Perceiving Emotions

The perception of the emotional information is divided in emotional information processing and emotional information seeking.

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 4.5).

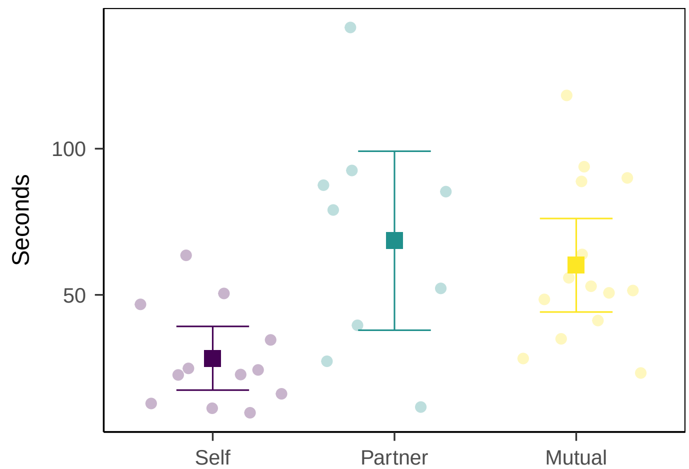


Figure 4.5: Total time (in seconds) spent looking at the perceiving-monitoring zone of the interface. Bars represent 95% confidence intervals.

An overall effect of the experimental condition on the time spent processing emotional information could be observed ($F(2, 32) = 6.24$, $MSE = 809.30$, $p = .005$, $\hat{\eta}^2_G = .281$ in a one-way ANOVA). Pairwise comparisons, depicted in Table 4.4, 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.

Table 4.4: Pairwise comparisons of the thee interfaces with respect to total time spent looking at the perceiving-monitoring zone of the EAT (p-values are adjusted with the Tukey method).

contrast	estimate	SE	df	t.ratio	p.value	cohens.d
Self - Partner	-40.20	12.54	32	-3.20	.008	-1.41
Self - Mutual	-31.81	11.19	32	-2.84	.021	-1.12
Partner - Mutual	8.39	12.15	32	0.69	.771	0.30

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 4.6).

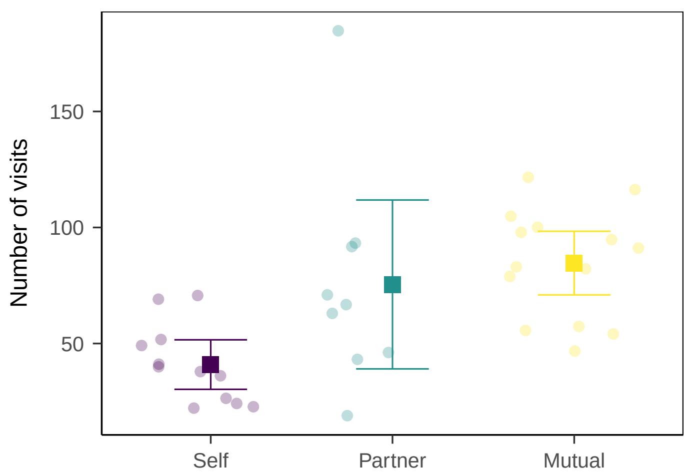


Figure 4.6: Total visits count in the perceiving-monitoring zone of the interface. Bars represent 95% confidence intervals.

An overall effect of the interface adopted on emotional information seeking could be detected ($F(2, 32) = 7.42$, $MSE = 886.76$, $p = .002$, $\hat{\eta}^2_G = .317$ in a one-way ANOVA). Pairwise comparisons, depicted in Table 4.5, 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.

Table 4.5: 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).

contrast	estimate	SE	df	t.ratio	p.value	cohens.d
Self - Partner	-34.53	13.13	32	-2.63	.034	-1.16
Self - Mutual	-43.73	11.71	32	-3.73	.002	-1.47
Partner - Mutual	-9.20	12.72	32	-0.72	.752	-0.31

4.5 Discussion

The planned analyses aimed at investigating whether a different use of, and access to, emotional information determine differences in use of an EAT during a computer-mediated collaborative task, which – beside evident shortcomings in the *collaborative* nature – may be considered representative of a CSCL application in distance learning. The reduced sample size upon which the analyses are based requires caution in interpreting the obtained results. The inter-individual differences in all measured 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 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. 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.

4.5.1 Expressing Emotions May not Depend Excusively on Social Sharing

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 *Partner-Oriented* and *Mutual-Modeling* condition, because of the absence of social sharing. 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.

Results failed to corroborate any of these assumptions given that neither an overall effect, nor differences in the comparison between each condition could be detected. The effect size to yield significant results was already ambitious with the planned $N = 48$ sample size, and was therefore even further undercut by post-hoc exclusions, for which even observed effect sizes above $d = 0.5$ (*Self-Centered* vs *Partner-Oriented* and *Self-Centered* vs *Mutual Modeling*) cannot yield a discernible difference. Hypothesis *H1* must therefore be provisionally rejected, even though the observed directions and size of the effects are, in part at lest, consistent with an increased expression of the emotions with a more social-oriented interface.

Notwithstanding the limits of the sample, it is also worth reversing the perspective and, rather, 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 of task. 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 support for *not* ruling out the interest of *intra-personal* interest in expressing emotions: the presence of an EAT could indeed serve as an *individual affordance* for emotional expression as a support for (implicit) emotion regulation. 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-in* idle time between enigmas or part of the task within each enigma, rather than for an urge to express and regulate their emotions. A more fine-grained analysis of the time of expression should therefore be taken into account.

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. Nevertheless, the experimental settings elicited considerable variation in the number of emotions expressed: with an increased sample size, the experimental plan could potentially contribute to assess the matter more thoroughly, for instance using a planned equivalence test with the aim to rule out differences, rather than detecting ones (Fidler, Thorn, Barnett, Kambouris, & Kruger, 2018; Lakens, 2017).

4.5.2 Emotional Seeking and Processing Seem Related to Social Sharing

In the second hypothesis, it has been posited that learners would seek and process the emotional information available on screen depending on the source and the comparison it facilitates. 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 $\hat{\eta}_G^2 = .317$, and of $\hat{\eta}_G^2 = .281$ for information processing. In both cases, thus, the experimental condition seems to account for a considerable amount of variation in the perceiving-monitoring function of emotional awareness.

On a more fine-grained level, though, 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* (Cohen's $d > 1$ for both measures) and *Self-Centered* vs. *Mutual-Modeling* (again Cohen's $d > 1$ for both measures), but not between *Partner-Oriented* vs. *Mutual-Modeling*, which was actually a more severe test (Mayo, 2018).

The lack of a detectable difference between the *Self-Centered* and the more social-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. Nonetheless, taking the raw measures as benchmark, it is reassuring to observe that in a task of 20 minutes, the time spent looking at emotional information is around 1 minute when information about the partner is available, compared to 30 seconds when it is not. As it is the case for the expression of emotions, though, information processing and seeking in the *Self-Centered* measures must be considered as support for an intra-personal interest for the use of an EAT, even in the absence of communication with the partner.

The comparison between the *Partner-Oriented* and *Mutual-Modeling* interfaces has deeper implications with respect to the *raison d'être* of an EAT. The lack of a discernible effect between the two social-oriented conditions 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 for 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 therefore also be assessed with the aid of self-reported measures about the perceived usefulness of direct and persistent comparison of learners' emotional states.

4.6 Post-Hoc Corollary Analyses

In this section, I provide the results of additional analysis that have not been planned before the study. First, I extended the analysis of eye-tracking measures using transitions between Areas Of Interest as an interesting measure of the use of an EAT. Second, I provide indications of the use of the EAT in real-time with respect to the appraisals and subjective feeling measures collected through the task. And finally, I take advantage of the use of the same task as in Fritz (2015) to conduct a small, internal meta-analysis that can be of interest for the use of the same task in future studies.

4.6.1 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, real-time phenomena in the overall thesis, it is worth investigating also transitions between the three main Areas of Interest (AOI) of the experimental task, 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, and thus contribute to better assess the perceiving-monitoring function of an EAT.

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 a keyboard. Therefore some transitions may have been lost due to the fact that each gaze-path may have been interrupted by a *detour* to the keyboard. Nevertheless, it is safe to assume that participants 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 lacks external validity and should 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 regardless of the group: 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 4.7 reports the number of transitions stratified by experimental condition between the 6 AOI organized in three rows such as each row displays the transitions between the same two AOI in both directions.

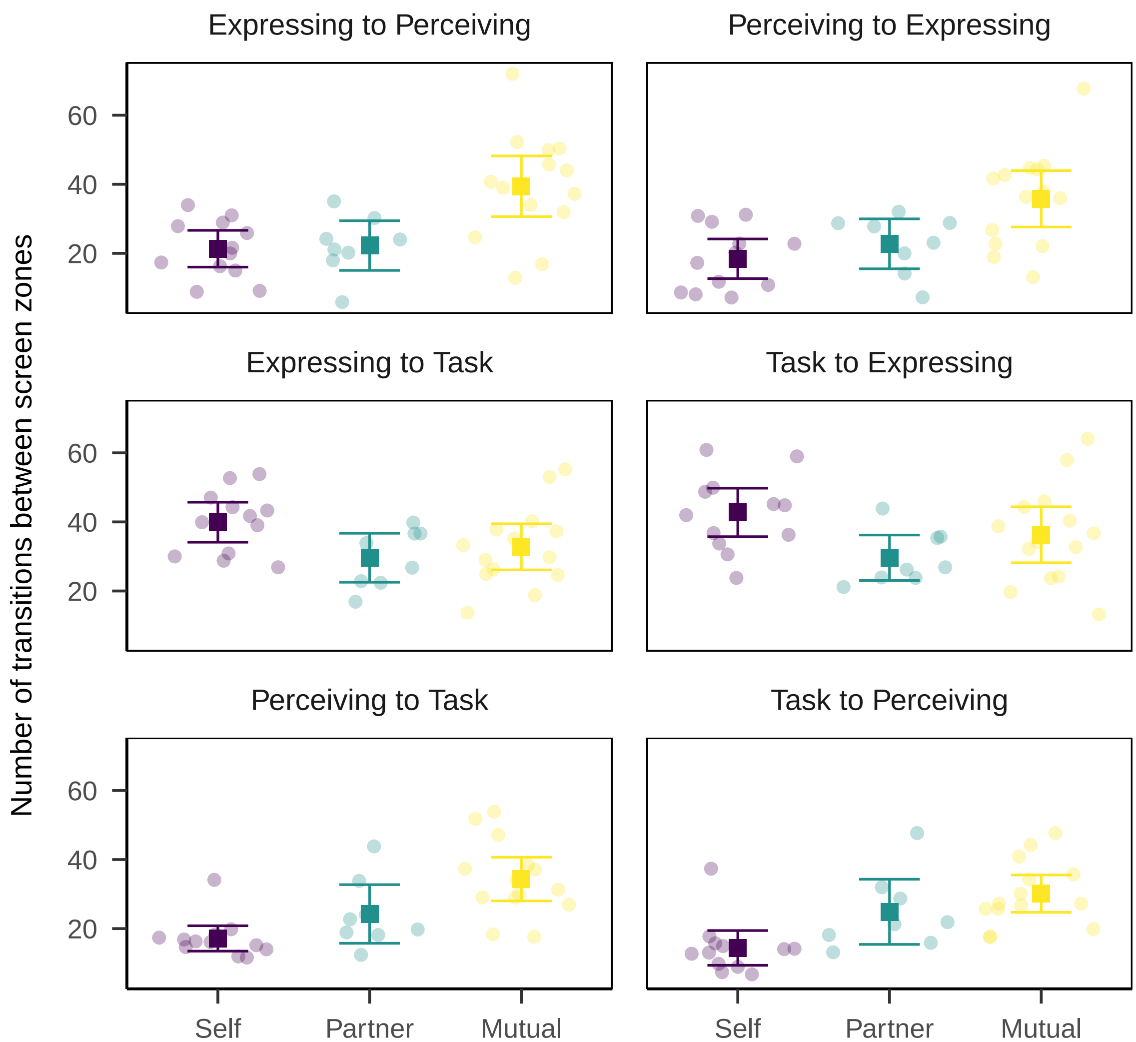


Figure 4.7: Number of transitions between Areas of Interest (AOI) on the interface. Transitions aggregated for $N = 34$ participants. Bars represent 95% confidence intervals.

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 *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.

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 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 confirmation is not available in the *Partner-Oriented* condition.

Finally, the *Partner-Oriented* condition seems once again *stuck in the middle*, and results for this group are difficult to assess due to the greater inter-individual variance that is also present in the other 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 (Bates, Mächler, Bolker, & Walker, 2014; Kuznetsova, Brockhoff, & Christensen, 2017; West, Welch, Gałecki, & Gillespie, 2015), was fitted to the data at hand in the following way: the aggregated 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 (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 4.6 using Kenward-Roger approximation for computing the p-value (Luke, 2017).

Table 4.7 reports the pairwise comparisons between the three experimental conditions stratified by the bi-directional path of the transition. Detectable differences in the pairwise comparisons can be observed between *Self-Centered* vs *Mutual-Modeling* and *Partner-Oriented* vs *Mutual-Modeling* in the transitions between expressing-

Table 4.6: 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

displaying and perceiving-monitoring. In the four comparisons, the more *social-oriented* interface obtained more transitions, in both directions, than the less *social-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 the path going from the task to the expression-displaying zone yielded a detectable difference 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, detectable differences were observed 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.

All things considering, transitions may represent a more interesting measure of the perceiving-monitoring function of emotional awareness compared to the information seeking and processing measures 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 dual-task nature of emotional 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.

Table 4.7: 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

4.6.2 Emotions and Time: Evaluating the Purpose of Real-Time Awareness

One of the main tenets of the present contribution is the advantage of *real-time* emotional awareness – at least to the extent that participants do not have to wait predefined stops to share their emotions. Some exploratory analyses on the sample were performed in order to check to what extent the *real-time* 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 4.8 shows all the $N = 483$ emotions that have been expressed by all the participants over the 20 minutes of the task, stratified by condition. For each observation, the value of Valence and Control/Power are displayed with respect to the elapsed time in the task. A Locally Estimated Scatterplot Smoothing (LOESS) – that is, non-parametric curve that best fit the empirical data (Jacoby, 2000) – is superposed to the raw data.

The graphic indicates that the smoother remain, overall, close to the neutral point, since data-points are evenly spread over the elapsed time in the task. It is worth noting, though, that in the *Mutual-Modeling* condition, both appraisals decreased as the task went on.

Subjective Feelings Over Time

The same analysis can be conducted with respect to the expression of subjective feelings over time. Figure 4.9 below plots the evolution of the expression of the 20 subjective feelings – which are part of the underlying affective space used in the study – aggregated 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

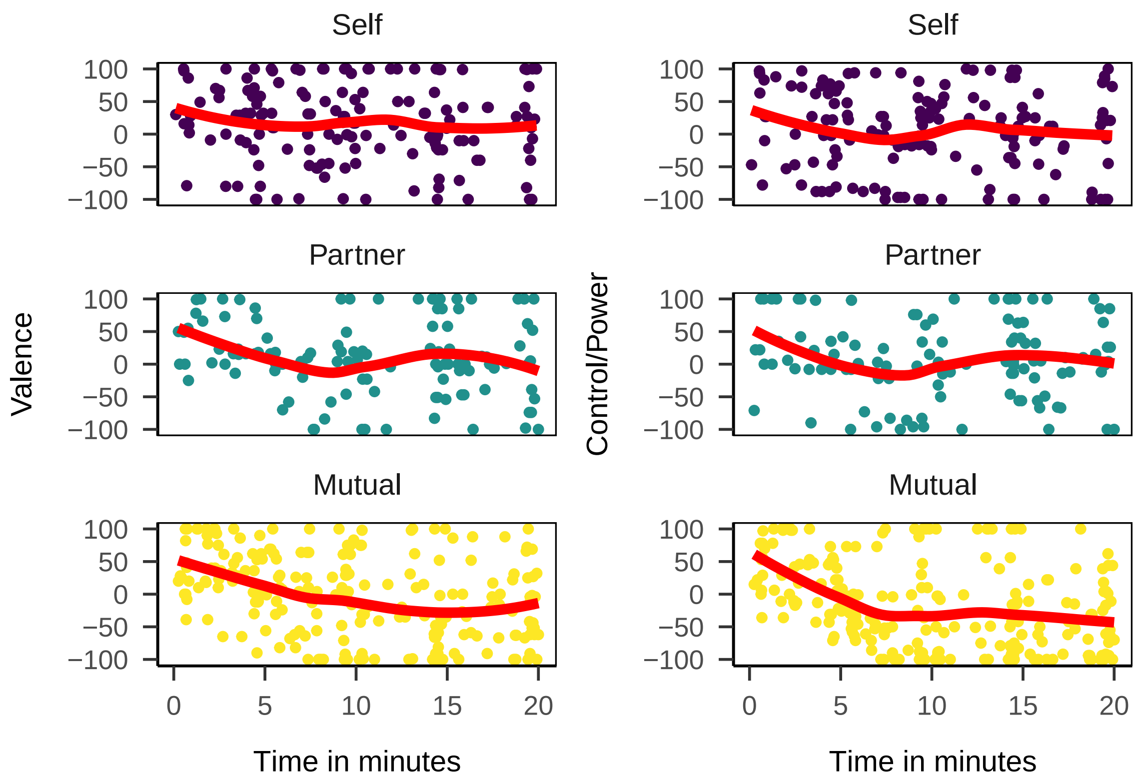


Figure 4.8: Evolution of the appraisal dimensions over time with a LOESS smoother ($N = 35$, all emotions of participants are aggregated per condition).

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 *Emphatic* 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.

4.6.3 Internal Meta-Analysis on Task Indicators

Taking advantage of the fact that the same task was adopted, under similar conditions, of a previous study (Fritz, 2015), an internal meta-analysis was performed on

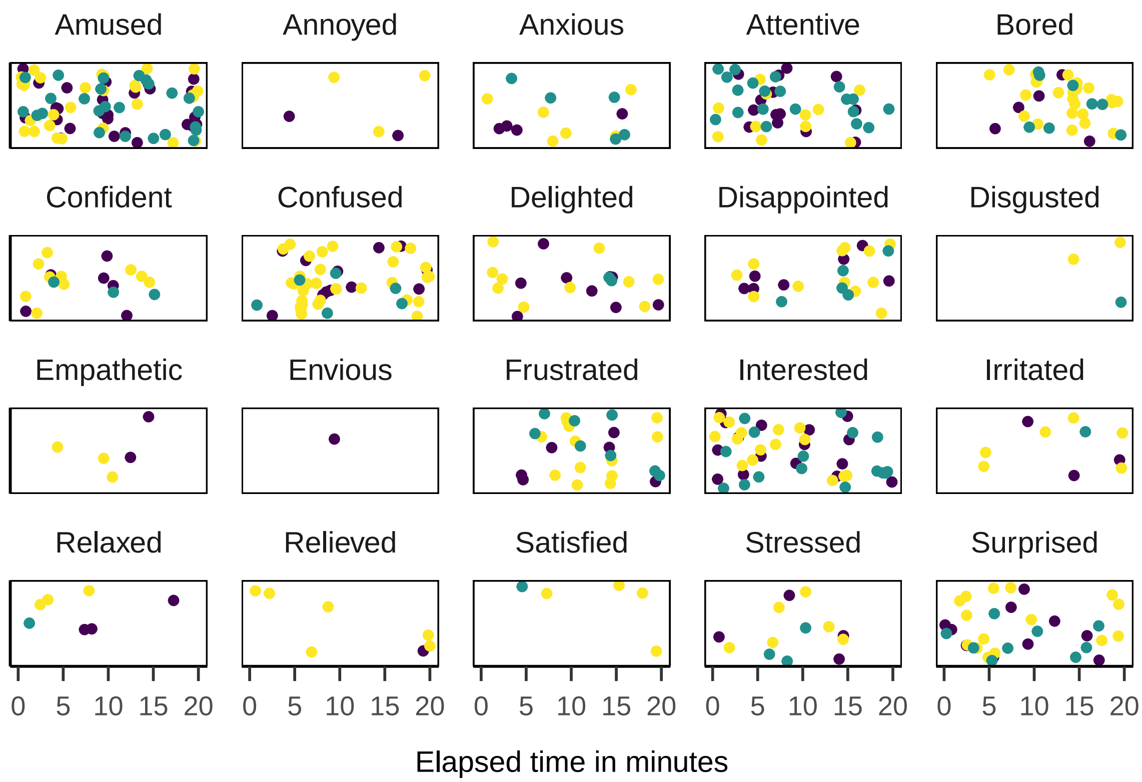


Figure 4.9: Expression of the subjective feelings. $N = 458$ emotions (out of 483) whose subjective feeling belongs to the underlying affective space used by the DEW, aggregated for the $N = 35$ participants. (Legend omitted for reducing space, see previous graphs.)

the point-estimate means for the three dependent variables adopted in the current contribution. The interest of the meta-analyses is two-fold. On the one hand, they provide a better assessment of the point estimates about the performance-based indicators of the use of the EAT. On the other hand, those same indicators will be used as references in a subsequent study in this contribution.

Each meta-analysis has been conducted using the R meta package, adopting the inverse of the variance weighting mechanism to account for differences in the sample size of the two internal studies. 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 the whole sample size of 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 $N = 16$ in Fritz (2015) and of $N = 35$ in Perrier (2017). Results, depicted in Figure 4.10, assess an estimated mean of 14.82 [13.14; 16.50] expressed emotions for the fixed effect model, and of 16.04 [11.16; 20.92] for the random effect model. The meta-analysis highlights the presence of considerable heterogeneity in the expression of emotions ($\tau^2 = 10.29$; $\tau = 3.21$; Iˆ2 = 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), whereas in Perrier (2017) they did not receive any guidance. 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.

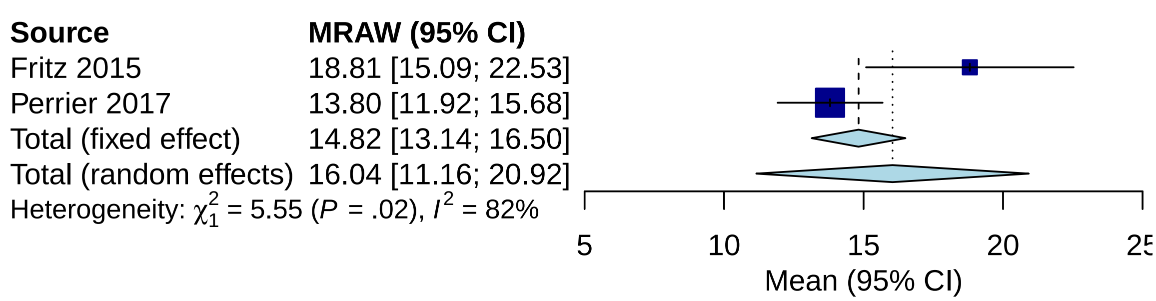


Figure 4.10: Internal meta-analysis of the number of emotions expressed in the experimental task.

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 ($N = 14$) in Fritz (2015), and only participants in the *Partner-Oriented* and *Mutual-Modeling* conditions ($N = 23$) in Perrier (2017), for the interface in these situations is identical or at least very similar with respect to the *social* information shared. Results, depicted in Figure 4.11, assess an estimated mean of 57.23 [47.51; 66.95] total visit duration, in seconds, for the fixed effect model, and of 56.99 [43.83; 70.15] for the random effect model.

The meta-analysis does not detect heterogeneity between studies (τ^2 = 40.81; τ = 6.39; Iˆ2 = 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.

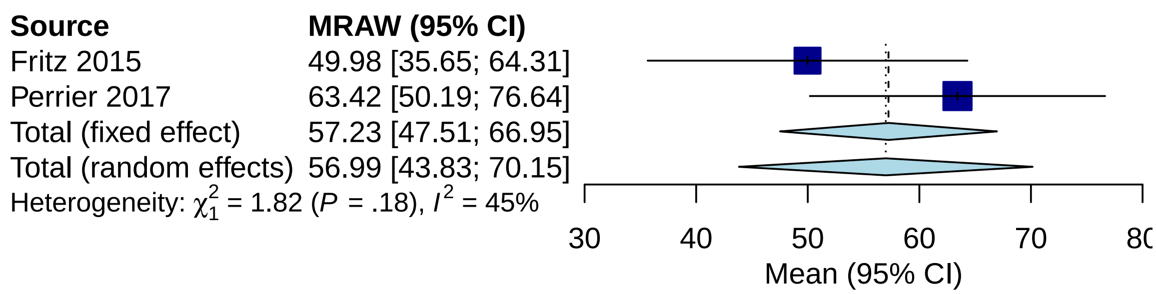


Figure 4.11: Internal meta-analysis of the time spent at processing emotional information available on screen.

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 = 14 in Fritz (2015) and N = 23 in Perrier (2017). Results, depicted in Figure 4.12, assess an estimated mean of 77.78 [66.71; 88.85] number of visits for the fixed effect model, and of 77.78 [66.71; 88.85] for the random effect model. The meta-analysis does not detect heterogeneity between studies (τ^2 = 0; τ = 0; Iˆ2 = 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.

4.7 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 following exclusion criteria, were randomly assigned to three different interfaces of the EAT – namely a *Self-Centered*, a *Partner-Oriented*,

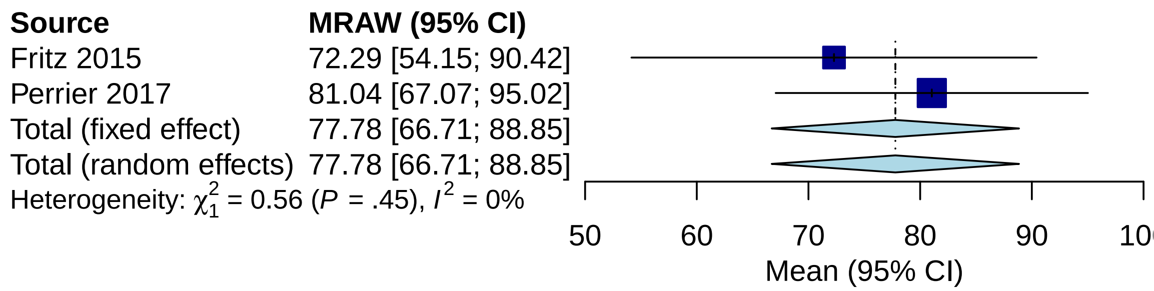


Figure 4.12: Internal meta-analysis of the number of times emotional information has been visited on screen.

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, since a detectable difference was not observed between the three conditions; and it 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 more than in the *Self-Centered* condition, but no detectable difference was observed 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. These indicators included the number of emotions expressed through the tool, the number of visits and seconds spent looking at the perceiving-monitoring part of the EAT, as well as the transitions between the more socially-oriented parts of the EAT interface. In all these cases, 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, Van Doorn, Heerdink, & Koning, 2011), these results seem to corroborate the usefulness of an EAT as a 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-expressing function of awareness. This fact seems to corroborate the usefulness of providing real-time 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, Lemaignan, Sangin, Nova, & Molinari, 2016).

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 (Matthew D. Lieberman et al., 2007; Matthew D. Lieberman, 2019; 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, DePaula, Dourish, & Sengers, 2007; Grandjean, Sander, & Scherer, 2008; Scherer, 2009).

4.7.1 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, Fischer, & Bos, 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. That could be obtained either by a direct replication of the current setting, but using pairs of participants, or by a split design in which part of the participants are randomly assigned to a simulated- or a real-partner. This second option would elicit a better and more reliable comparison, but its interest would be fairly limited to the *validation* of a simulated participant, which was an auxiliary instrument in the study. A direct replication with real collaboration, on the other hand, would investigate the subject matter more thoroughly, even though the comparison with present results would have to take the difference in time and setting into account.

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 real-time emotional awareness: whether and when participants would decide to display the EAT may convey pivotal evidence about the usefulness of real-time emotional awareness.

4.7.2 Aknowlegments

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

Chapter 5

Emotion Awareness in Asynchronous and Individual Settings

The study illustrated in the previous chapter suggests the presence of a genuine interest in the use of an Emotion Awareness Tool (EAT) in a real-time, computer-mediated collaborative task. Nevertheless, this does not exclude the possibility that emotional awareness may be beneficial above and beyond these conditions, for instance in asynchronous and individual conditions. It is in fact widely accepted that one of the main drawbacks in remote learning is the lack of social presence: learners often feel alone, without any contact with their colleagues (*e.g.*, Jacquinot, 1993; Paquelin, 2011; Parkinson, 2008; Sung & Mayer, 2012). An interest towards an EAT may therefore be motivated by the fact that emotions are used as a *proxy*, a strong reminder that there is someone else *out there*, even without the need to integrate the emotional information as instrumental to task at hand (Buder, 2011; Dourish & Bellotti, 1992). In this regard, the use of an EAT may contribute to provide learners with a shared understanding of the affective implications of the distance learning conditions. For instance, sharing emotions with their colleagues may, one the one hand, provide useful information as social reference, on the basis of which learners can regulate their own emotions in facing the challenges of remote learning conditions (A. H. Fischer & Manstead, 2016; Gross, 2014; Gross, 2015; Winne, 2015). On the other hand, the affective dimensions created through the EAT may contribute to foster a sense of belonging and cohesion to the group, even in the absence of face-to-face and/or synchronous exchanges (S. Barsade, 2002; S. G. Barsade & Knight, 2015; Parkinson, 2008, 2011; Rimé, 2005, 2009; Salas, Grossman, Hughes, & Coultas, 2015; Van Kleef & Fischer, 2015; Van Kleef, Heerdink, & Homan, 2017).

This second empirical study will investigate whether real-time collaboration represent a necessary condition for the usefulness of an EAT in distance learning, or whether the social and relational information it conveys can be useful even in an asynchronous and individual condition. In this regard, students of two classes following the same course in a blended Master in Learning and Teaching Technologies will be provided with the possibility to use an EAT during the periods of remote learning, through which they can share their emotional states with all the other members of the class. Their use of the tool and their attitude towards the usefulness of it will be monitored at different time during the semester in order to assess the usefulness of the EAT. In this intent, the chapter also introduces the Emotion Awareness Usefulness (EAU) survey, which is derived from dimensions described in the literature, and aims at investigating to what extent the presence of an EAT is instrumental in understanding, interpreting, and experiencing emotion (Boehner, DePaula, Dourish, & Sengers, 2007).

The end of the chapter provides corollary analyses as a preliminary attempt in investigating whether the adoption, use and perception of the EAT are related to learners' emotional competence (Brackett, Bailey, Hoffmann, & Simmons, 2019; Brackett, Rivers, Reyes, & Salovey, 2012; Scherer, 2007; Schlegel & Mortillaro, 2018). A link between three sub-competences – namely the appraisal, the communication, and the regulation competences (Scherer, 2007) – and the use of an EAT is also proposed.

By implementing the use of an EAT in longitudinal, collective, and ecological settings, the present chapter aims at investigating the adoption, use and perception of an EAT by learners in remote learning environment. The present chapter may therefore be of interest to the broader field of distance learning and, more specifically, to the investigation of the relationship between affective states and technology enhanced learning (Arguedas, Xhafa, Daradoumis, & Caballe, 2015; R. S. J. D. Baker, D& apos;Mello, Rodrigo, & Graesser, 2010; D'Mello, 2013; D'Mello & Graesser, 2012; Lehman, D'Mello, & Graesser, 2012; R. C. D. Reis et al., 2015; Rachel Carlos Duque Reis et al., 2018). The chapter also investigates emotions at the group level (*e.g.*, S. Barsade, 2002; Cheshin, Rafaeli, & Bos, 2011; Keltner & Haidt, 1999; Parkinson, Fischer, & Manstead, 2005; Smith, Seger, & Mackie, 2007; Van Kleef & Fischer, 2015; Van Kleef, Heerdink, & Homan, 2017) and may therefore be of interest in studies using groups as unit of observation.

5.1 Study Overview

At the origin, awareness in Computer-Supported Cooperative Working (CSCW) was limited to cues about the presence of other co-workers: whether they were online and on what they were working on (Grudin, 1994; Gutwin & Greenberg, 2002; Gutwin, Greenberg, & Roseman, 1996). Progressively, awareness has assumed a broader perspective and, especially in computer-mediated collaboration and Computer-Supported Collaborative Learning (CSCL), there is nowadays a consensus about the need to provide – through awareness tools – information about others, which is inherently linked to social and relational phenomena (Arguedas, Xhafa, Daradoumis, & Caballe, 2015; Buder, 2011; Janssen & Bodemer, 2013; Janssen, Erkens, & Kirschner, 2011). An Emotion Awareness Tool (EAT) is deeply rooted in this perspective, since the information shared is not directly part of the collaborative task. In other words, using Janssen and Bodemer (2013) division between the *content* and the *relation* spaces already mentioned also in the previous chapter, an EAT may be limited to the *relational* space if learners do not make the effort to integrate that information also in the *content* space. This fact may therefore limit the usefulness of Emotional Awareness to enhancing the social presence. That is, emotions shared through the EAT are used as a *proxy*, a strong reminder of the presence of other learners, but without integrating the emotional information into the task at hand. This phenomenon would not be inherently bad, since it is widely accepted that one of the main drawbacks in distance learning is the sentiment of loneliness and isolation (*e.g.*, Carswell, Thomas, Petre, Price, & Richards, 2000; Conrad, 2002; Jacquinot, 1993), but it would question whether the result is worth the effort: social presence may be sustained with cues that are closer to the *content space* compared to the dual-task (Pashler, 1994) imposed by the use of an EAT while performing the learning activities. Performance-based indicators of interest in emotional information sharing, seeking and processing – obtained in the randomized trial with three different interfaces illustrated in the previous chapter – seem to corroborate a genuine interest in emotional awareness as instrumental to the task at hand. That does not exclude the possibility, though, of an interest in using the EAT exclusively as a form of social presence. Therefore, it is worth investigating the usefulness of an EAT in an asynchronous and individual situation: can it still be useful?

In order to investigate the matter, the adoption of an experimental approach as the one used in the previous study is nevertheless inadequate for at least the following reasons. First, using a task of a relatively short time may fail to produce a sense of belonging to a group, especially in a computer-mediated settings. It is therefore more reliable to measure the usefulness of the EAT in an extended period of time, ideally in order to assess whether the perception of usefulness evolves in time. Second, if collaboration is removed, participants would be sharing emotions with, and seeing emotional episodes of an estranged person, without any connection in time, space or purpose. Thus, the social presence should rather be elicited in a group that has already a *raison d'être*, and whose members share as many elements as possible, even though they are not directly collaborating. Third, in an asynchronous and non collaborative situation, emotions may be elicited by a wide range of elements which are more difficult to pinpoint to something that has happened synchronously. It is thus important that the context in which emotional episodes emerge is as ecological as possible, for them to be representative as a *proxy* of the person.

For these reasons, the present study adopts a longitudinal plan (Fitzmaurice, Laird, & Ware, 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, during which students are often assigned a small project to submit before the following period begins. 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, a mechanism comparable to the Experience Sample Method [ESM; Csikszentmihalyi & Larson (2014)] or the Ecological Moment Assessment [EMA; Shiffman, Stone, & Hufford (2008)] already used in distance learning situations (*e.g.*, Molinari, Trannois, Tabard, & Lavoué (2016), but see for instance Scollon, Kim-Prieto, & Scollon (2003) for a critical assessment of the method). 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).

Compared to the previous study, the emotional information of more than two students will be shown on the perceiving-monitoring part of the interface (Buder, 2011; K. Schmidt, 2002). As a consequence, the way emotional awareness is graphically represented on screen must also be adapted in order to convey a *grouped* representation of a whole class (Berset, 2018; Fritz, 2016b). This technical modification has consequences on a more abstract and theoretical level, since *individual* or *dyadic* emotions are replaced by *group* emotions (*e.g.*, S. Barsade, 2002; Keltner & Haidt, 1999; Parkinson, Fischer, & Manstead, 2005; Smith, Seger, & Mackie, 2007; Van Kleef & Fischer, 2015; Van Kleef, Heerdink, & Homan, 2017). The use of the EAT may therefore be influenced by collective dynamics. For instance, Cheshin, Rafaeli and Bos (2011) found supporting evidence that emotional contagion (S. Barsade, 2002; Hatfield, Cacioppo, & Rapson, 1993) can also happen in virtual teams, where computer-mediated communication is exclusively text-based.

In this regard, comparing two classes of the same course in different years may therefore also contribute to investigate to what extent the adoption, use and perception of the usefulness of an EAT is determined *mainly* by individual characteristics or is also influenced by *interactions* between the individual and the group (Boehner, DePaula, Dourish, & Sengers, 2007; Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016). Even though students would not be randomly assigned to the classes, it is difficult to figure out systematic factors at play, which would determine why particular students would apply in a specific year rather than another, especially in a Master that has always been very heterogeneous in the students' profile.

Finally, whereas the number of emotions expressed can be maintained as an indicator of the use of the EAT, the eye-tracking measures used to detect emotional information seeking and processing cannot be replicated in a longitudinal plan. For this reason, the study also implements a tentative survey aiming at measuring the perceived Emotion Awareness Usefulness (EAU). The survey – thoroughly depicted in the material section – combines 7 dimensions retrieved from the literature:

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, 2013), 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 where they take place (Norman, 2013). The presence of an EAT may prompt users to 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 (*e.g.*, Gunawardena & Zittle, 1997; Tu & McIsaac, 2002). 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, Scherer, Grandjean, & Sander, 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.
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 and consequences on their behavior (*e.g.*, Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016; Van Kleef, 2010, 2018; Van Kleef, Heerdink, & Homan, 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, Ainsworth, & Crook, 2012; Molinari, Chanel, Bétrancourt, Pun, & Bozelle, 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, Daradoumis, & Xhafa, 2016; Gross, 2014; Gross, 2015; Järvenoja, Volet, & Järvelä, 2013). The presence of the EAT, both in terms of learners own emotions and that of the colleagues, may facilitate regulatory processes.

5.2 Research Questions

The present study aims at investigating four main topics. The first consists in the assessment of whether synchronous collaboration is a *necessary* condition for emotional awareness to be perceived as an integrated information into the task. If an EAT is perceived useful even when participants are not directly collaborating, it will rather suggest that emotion awareness is a *sufficient* condition for enacting social presence of others, regardless of the collaborative or individual nature of the task (Mackie, 1965). In other words, emotion awareness could act as a *proxy* for students not to feel alone during distance learning and provide social reference as to how colleagues are feelings (S. G. Barsade & Knight, 2015; Hareli, Moran-Amir, David, & Hess, 2013; Van Kleef, Heerdink, & Homan, 2017), as well as enhancing group belonging and cohesion during remote learning (Anderson, Keltner, & John, 2003; Keltner & Haidt, 1999; Salas, Grossman, Hughes, & Coultas, 2015; Van Kleef & Fischer, 2015). The first research questions (*Q1*) thus investigates what is the use and the perception of emotion awareness usefulness of students in a remote, individual learning environment, and whether the use or the perception of usefulness change over time.

The second purpose of the study is to contribute to assess whether the use of an EAT depends *mainly* on the individual characteristics of the students or is *also* determined by the interaction between students using the tool at the same time (Dillenbourg, Lemaignan, Sangin, Nova, & Molinari, 2016). Taking advantage of the fact that two classes will use the EAT under *more or less* equivalent external conditions (same course, same program, same learning environment, ...) – and assuming no systematic factor at play determining particular type of students in one class compared to the other – differences in the use or perception of the tool between one class and the other should corroborate the effect of interaction rather than individual characteristics alone. The second research question (*Q2*) therefore investigates whether differences in the use or the perception of the EAT can be detected between one class and the other.

The third aim of the study is to assess the perceived usability (Tullis & Albert, 2013) of an EAT in a longitudinal and asynchronous context. The previous assessment of the tool was conducted in a usability test, with the same fixed time and simulated collaborative task as in the previous chapter (Fritz, 2014) and is therefore worth investigating what is the perceived usability of the EAT in an ecological context.

The forth and last purpose of the study is to assess to what extent the moment-to-moment emotions expressed by learners are then recollected correctly after a period of time, both with respect to their own expressed emotions and that of their colleagues. There is in fact evidence in the literature suggesting that emotional episodes have a privileged access to memory (Brosch, Scherer, Grandjean, & Sander, 2013; Kensinger & Ford, 2020; Montagrin, Brosch, & Sander, 2013; Pool, Brosch, Delplanque, & Sander, 2015; Rimé, 2005), and it is therefore worth exploring whether the presence of an EAT contributes in somehow in *anchoring* the individual and collective affective experience during distance learning.

No hypothesis is nevertheless posited for any research question, because the setting is too unpredictable. The EAT has never been used before in longitudinal studies, which are by nature more prone to unforeseen events. For instance, it is not even warranted that the EAT will be used in the first place. Furthermore, the measure of emotion awareness usefulness is tentative and non-validated. Consequently, positing hypothesis without pre-registering them would expose any finding to well-founded doubts of questionable research practices (John, Loewenstein, & Prelec, 2012; Makel, Hodges, Cook, & Plucker, 2019), especially Hypothesizing After the Results are Known (HARKing; Kerr (1998)). All research questions should therefore be interpreted as non-confirmatory (Scheel, Tiokhin, Isager, & Lakens, 2020).

5.3 Methods

I report how I determined the sample size, all data exclusions (if any), all manipulations, and all measures in the study.

5.3.1 Participants

33 students (22 women, 6 men, and 5 missing values) 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 missing values). Students belonged to two classes that took the one-semester course in two successive years during the period of the thesis (2015-2020): 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 even more 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 differ with the other in the way social links could have been created, not disposing of the same weeks with on-site courses.

The use of the EAT was warranted as a pedagogical activity, since the use of technological tools in learning situations is an integral part of the Master's program. Given no participant can be forced to take part to a study, though, 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. 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 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.

5.3.2 Material

Configuration of the Emotion Awareness Tool

The Dynamic Emotion Wheel Research Toolbox (DEW-RT) was adopted for both cohorts. The configuration of the Dynamic Emotion Wheel (DEW), depicted in Figure 5.1, is implemented as follows.

For the expressing-displaying function of awareness, the same EATMINT circumplex used in the study depicted in the previous chapter and also in Fritz (2015) was adopted. As a reminder, this affective space comprises 20 emotions organized over the two appraisal dimensions *Valence*, prompted with the question *is the situation pleasant?*, and *Control/Power*, prompted with the question *is the situation under your control?*. 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 the chapter about the DEW-RT, within the section dedicated to the EATMINT affective space.

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 in a specific context, such a text document or a categorization, 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 feelings expressed by the participant herself;
2. A *Partners-Oriented* word cloud, depicting the last 100 feelings expressed by

the other members of the class, that is all the feelings bar that of the participant herself;

3. A *Collective-Oriented* word cloud, depicting all the emotions expressed by the whole class since the first use of the EAT, that is, the emotions of the participant herself, plus that of the other members.

Figure 5.1: Interface of the EAT for a participant, depicting the expressing-displaying and perceiving-monitoring parts of the tool.

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 present *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). All things considering, though, word clouds are relatively known graphical representation, 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. It is also warranted by the fact that, in the usability test conducted about the DEW (Fritz, 2015) adopting the very same interface than in the previous study, eye-tracking measures clearly showed that the subjective feelings in the emotion time line were sought more often than the line charts with the appraisals dimensions. 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 context, the temporal reference conveys limited information, especially considered 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. 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 in the DEW-RT is an open issue that has just started to be investigate (see for instance Berset, 2018).

Emotion Awareness Usefulness Rating

According to Boehner and colleagues (2007, p. 207), "success of [an affect-aware] system is measured by whether users find the system's responses useful for interpreting, reflecting on, and experiencing their emotions." Based upon this perspective, the study introduces a tentative scale that aims at measuring Emotion Awareness Usefulness (EAU) with respect to 7 dimensions identified in the literature (see also the study overview). The scale comprises 7 items expressing one of the 7 dimensions, for which participants expressed their accord on a scale from 1 (strongly disagree) to 10 (strongly agree) according to the item alone (*i.e.*, with no reference to the dimension):

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 only one per dimension. This is a potential shortcoming from a reliability and validity standpoint, but, especially in a repeated measure design, brevity is of essence. Once again, the convenience sample of the study provides some leverage on the formulation of items, since MALTT students are, for instance, familiar with terms such as *regulation*, which may be too technical for other populations and therefore require a more explicit formulation.

System Usability Scale

The System Usability Scale (SUS; Brooke (1996)) is a widely adopted scale that measures the usability of a tool. 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 are as follows:

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 concerned with emotions in a workplace, which has relevant congruence with interpersonal dynamics in remote learning. The test is divided in 4 sub-competences tests, 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 then choose among different natural language words the one that best describe the displayed emotion.

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 and must identify the two appropriate strategies vs. the two inappropriate ones.

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. 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 ranging from 0 (low competence) to 1 (high competence). The 4 scores can 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 have felt each of the 20 subjective feelings of the EATMINT circumplex ; and (2) their colleagues have 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.

5.3.3 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 my thesis and that data could be used for research. 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*. A scale from 1 to 10 allowed students to express their agreement which each dimension of the survey.

Demo Survey

Once filled the *Expectancy* survey, students were introduced to the EAT through a demo. They discovered the interface of the tool they will use 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 the EAT for Distance Periods

Towards the end of the course, 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 tool.

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 exercice.

Geneva Emotional Competence Questionnaire

In the private workspace 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 in addition to the questions of the test.

5.3.4 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.

5.4 Results

Results are based on $N = 30$ participants, having 3 students not filled both the *Halfway* and the *Final* survey, and were therefore excluded from the analysis considering they have dropped of the either from the course or the research. Table 5.1 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 5.1: 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

5.4.1 Expressing Emotions

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. The fact that those students had not expressed emotions, though, does not systematically rule out the fact that they have not used the tool at all: potentially, they could have logged in just to see the emotions of others. Technically, the DEW-RT registers the log for every access, but students were not informed beforehand of this possibility. Thus, data about accesses has not been extracted – and therefore not used – and the 4 students are kept in 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 5.2 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 5.3 The classes have similar profiles in the cumulative number of emotions expressed over time, alternating phases in which they express emotions with periods of pauses.

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 seems to be limited. It is nevertheless worth exploring whether the difference in expressing is also corroborated by differences in the perception of usefulness.

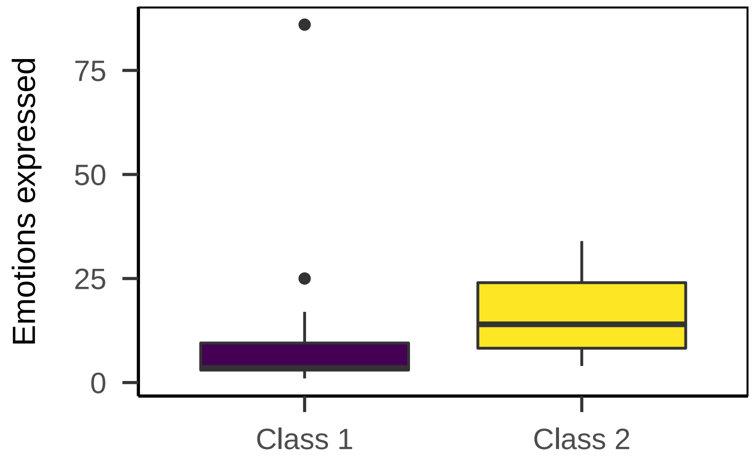


Figure 5.2: Boxplot comparing the expression of emotions between the two classes.

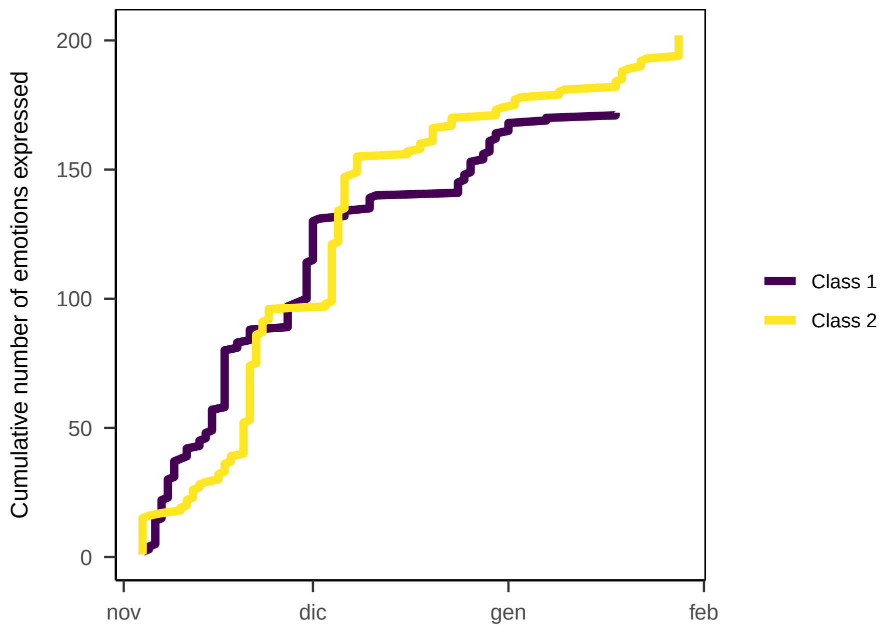


Figure 5.3: Cumulative number of emotions expressed through the EAT

5.4.2 Emotion Awareness Usefulness

Being the first adoption of the Emotion Awareness Usefulness (EAU) scale, it is worth beginning with some exploratory data analysis aiming at determining to what extent the self-reported rating over the 7 dimensions can contribute to assess (1) whether there is a change over time of the perceived EAU, and (2) whether the two classes differ in the perception of the EAU. In this regard, it is useful to start by evaluating the sensibility of the scale in terms of overall dispersion for each of the 7 underlying dimensions, namely *Frequency*, *Affordance*, *Social Presence*, *Self-Understanding*, *Understanding Others*, *Self-Other Comparison*, and *Self-Regulation*. In Figure 5.4 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 over the entire range of the rating, suggesting that each dimension has a good sensibility.

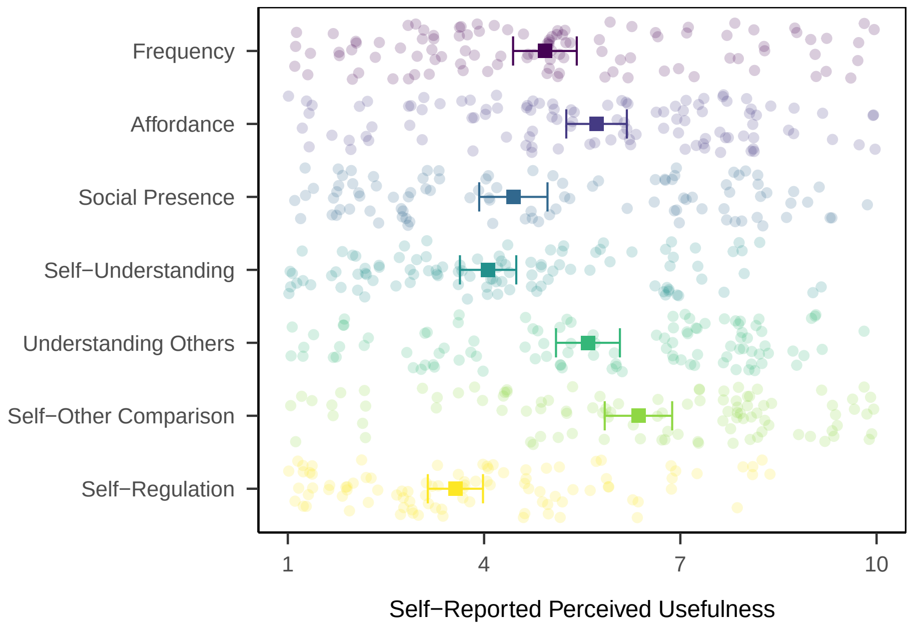


Figure 5.4: Overall ratings of each of the 7 EAU dimensions, $N = 798$ ratings. Bars represent 95% confidence intervals around the overall mean.

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 5.2 depicts means and standard deviations of the 7 dimensions stratified for both classes across the 4 longitudinal surveys.

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, Welch, Gałecki, & Gillespie, 2015) is fitted with the following components. 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

Table 5.2: 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)
Total	**4.23 (2.53)**	**4.34 (2.27)**	**2.66 (1.90)**	**2.92 (1.87)**

belongs to with 2 conditions (*Class 1* or *Class 2*). All two-way interactions between pairwise covariates, as well as the three-way interaction, are 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 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, Bolin, & Kelley (2019); McElreath (2020); Singmann & Kellen (2020); West, Welch, Gałecki, & Gillespie (2015) 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, Bolker, Westfall, Aust, & Ben-Shachar, 2020; Singmann & Kellen, 2020) R package. A Type III analysis of variance of the multilevel linear model detects 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, a the three-way interaction between group, longitudinal survey and EAU dimension was not detected. Results are depicted in Table 5.3 using Kenward-Roger approximation for computing the p-value (Luke, 2017).

Table 5.3: 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	27.996	0.557	.462
survey	3	716.778	63.477	<.001
dimension	6	714.003	29.403	<.001
group:survey	3	716.778	1.172	.319
group:dimension	6	714.003	3.769	.001
survey:dimension	18	714.003	2.085	.005
group:survey:dimension	18	714.003	0.664	.848

Results are graphically depicted in Figure 5.5, in which the EAU evaluation is stratified by the 7 dimensions of the scale and the 4 longitudinal surveys, as well as grouped by 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.

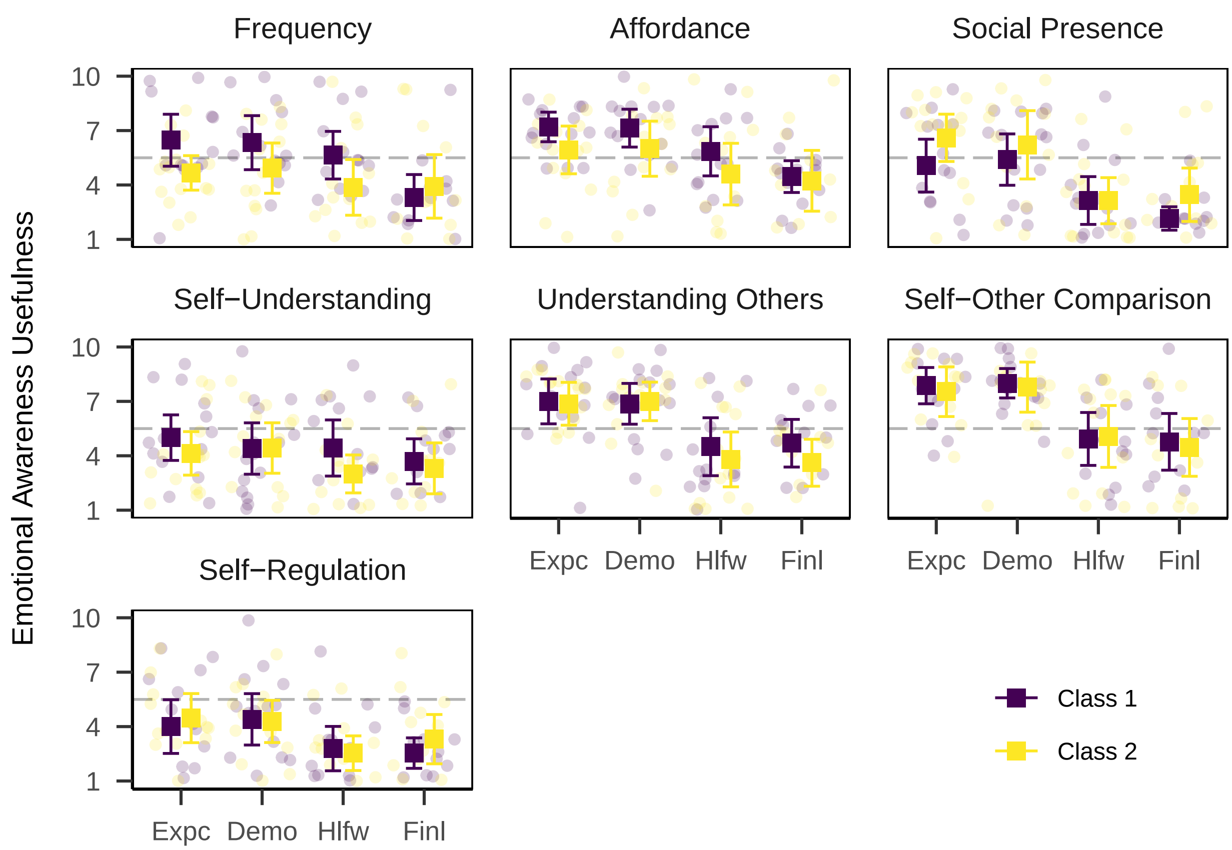


Figure 5.5: EAU rating over longitudinal surveys stratified by dimensions and grouped by class. Bars represents 95% confidence interval and the dashed gray line the median point of the scale.

Results of the post-hoc contrasts averaged over group and stratified across EAU dimensions are illustrated in Table 5.4. 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 decrease), followed by *Understanding Others* and *Affordance* (more than 2 rating points decrease), and finishing with *Frequency* and *Self-Regulation* (more than 1 rating point decrease).

5.4.3 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 5.6 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. Ad 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 compared to the other items, even if not as low as the first item about frequency.

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 &

Table 5.4: 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

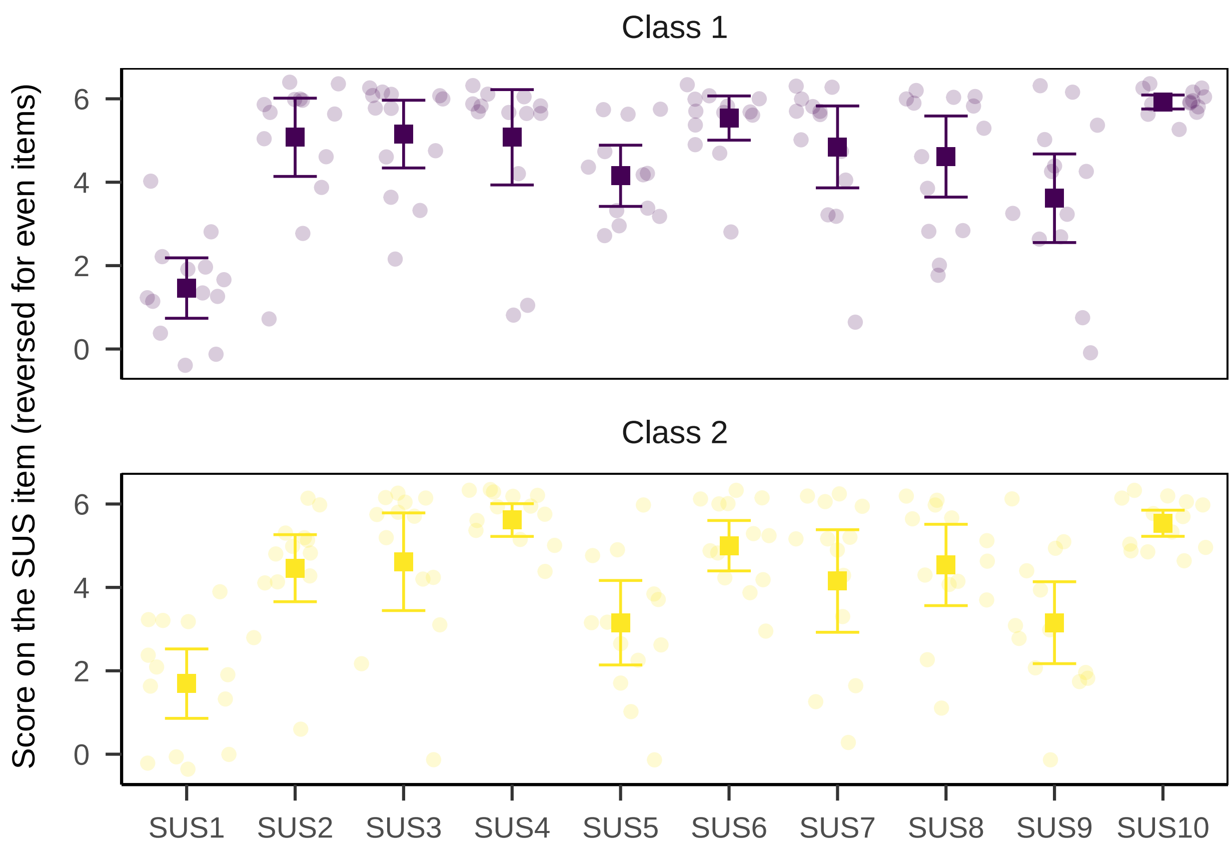


Figure 5.6: Rating of the individual items of the System Usability Scale (SUS) for N = 26 participants.

Sauro, 2018).

5.4.4 Recollection of Individual and Collective Emotions

In the *Final* survey, that is approximately two weeks after the end of the semester – and therefore the last day in which the EAT could be used by students during the STIC I course - participants were asked to recall the frequency with which (1) they have experienced the 20 subjective feelings of the underlying affective space as individuals, 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 . . . compares three relative frequencies of the 20 subjective feelings of the underlying affective space adopted by the EAT. The first frequency is empirically retrieved using the relative frequency of expression of each subjective feelings compared to all the expressed feelings. The frequencies are mapped so that the most expressed feelings (e.g. *Attentive* and *Interested*) represent the upper bound, and correspond to the *Very often* frequency, whereas the least expressed feelings (e.g. *Surprised* and *Disgusted*) represent the lower bound, and correspond to the *Never* frequency. The second depicted frequency consists in how often participants recalled their colleagues have experienced the specific feelings. It is worth reminding at this purpose that from the interface of the EAT it was not possible to infer which specific colleague expressed a particular emotion, and therefore it was potentially possible to retrieve the frequency from the *word clouds* of the interface. Finally, the third frequency is the self-reported frequency by which participants experienced each subjective feeling themselves. The class and the self frequencies were collected by asking participants to rate the frequency on the same 5-point scale used in the graph (*i.e.*, from never to very often).

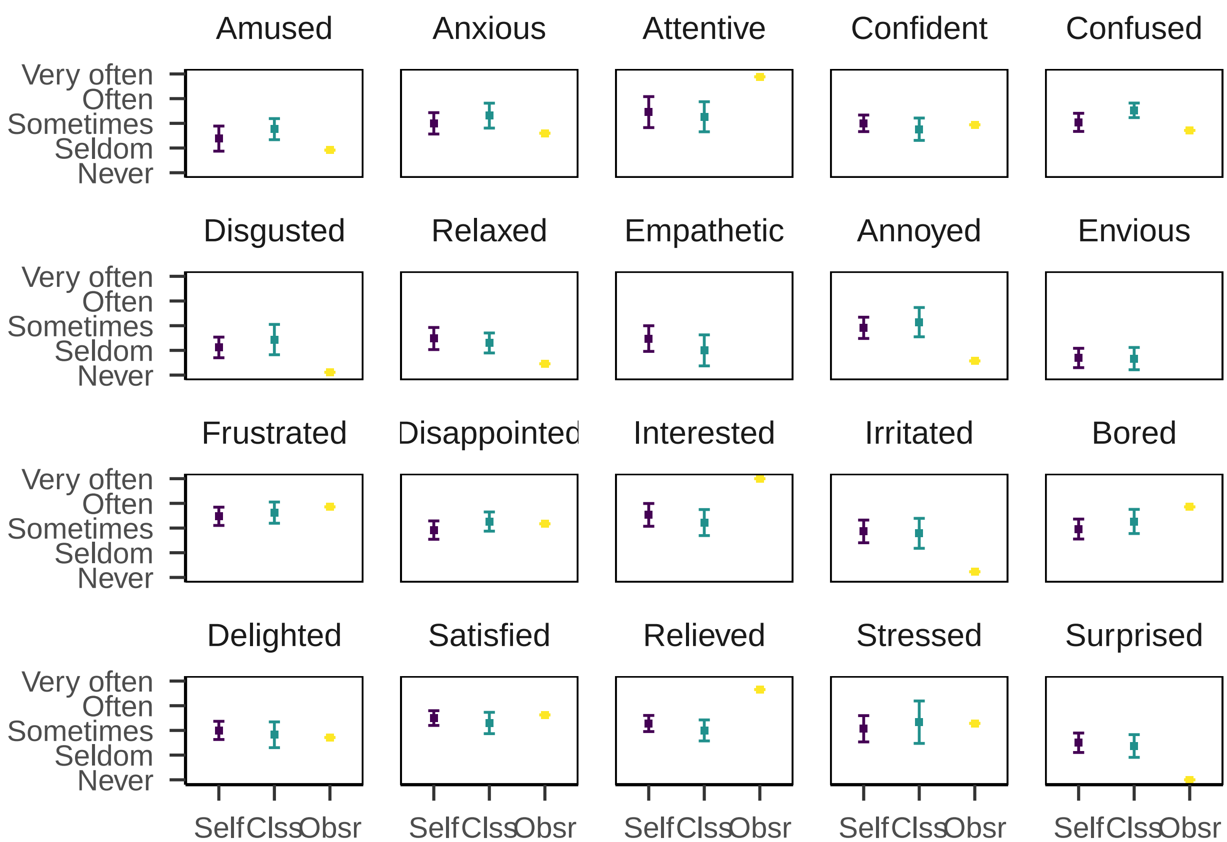


Figure 5.7: Comparison between the observed relative frequency of expressed feelings and the reported frequency recollected for the participants themselves, as well as the frequency attributed to the class as a whole.

5.5 Discussion

The overall picture emerging from the implementation of an EAT in asynchronous and individual settings can be roughly divided in two considerations. On the one hand, there is encouraging evidence about the expectations learners have towards the potential usefulness of an EAT. On the other hand, data highlight how these expectations are not sustained by the concrete use of the tool. Finally, the two classes showed consistency in both expectations and use, which is in itself a phenomenon worth discussing.

5.5.1 *Great Expectations*

As a reminder, participants to the study were enrolled in a blended course organized over three periods of time. During the first period, students had the occasion to get familiar with distance, technology enhanced learning. When they filled the *Expectancy* survey, they therefore already had a fresh, first-hand experience of the difficulties they had to face during the remote period. Expectations were similar between classes and will therefore be discussed on the aggregated score (on a 1-to-10 range). In this regard, 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. 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). 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. 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 difficult to foresee: oftentimes it is overestimated, as in this case ($M_{total} = 5.57$), since the actual use turned out to be very limited. Finally, the moderate expectations for *Social Presence* ($M_{total} = 5.83$), compared with higher expectations for the 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.

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. This is also consistent with the overall perceived usability ($M_{SUS} = NA$).On the hand, though, the tool failed to change their mind on the dimensions there were rated moderate to low.

5.5.2 Limited Impact of Emotion Awareness

The *great expectations* did not stand the (longitudinal) test of time, though. Overall, the EAT has been adopted only by a handful of users throughout the course, with a mean of expressed emotions below 13. This mean is also inflated by a few participants expressing most of the emotions. The limited use of the EAT is corroborated by the Emotion Awareness Usefulness (EAU) ratings, whose dimensions all decreased between the *Expectancy* and the *Final* survey, except for the *Self-Understanding* dimension, which was already low. The emotional information did not make any lasting impression on participants either, since self-reported recollection of subjective feelings' frequency of expression both on the individual and collective level was rated in a similar way, that is neutral on the scale from *Never* to *Very often*.

The answers to the open-ended questions proposed halfway and at the end of the semester mainly highlight that students forgot to use the tool, even when they have found it useful. On a superficial analysis, one may infer that the tool is not interesting enough, for it to be remembered. That is, the effort to recall its existence, open, and use it outweights the benefit derived from using it. The course as a whole, though, is very demanding in terms of different environments, tools, and documentation to be consulted.

5.5.3 Consistency Between Classes

5.6 Corollary Analysis

The chapter proposes two corollary type of analysis that may be of interest for future contributions. The first concerns the assessment of the Emotion Awareness Usefulness (EAU) scale in terms of its structure and reliability. The second explores a provisional link between emotional competence, using the scores obtained through the Geneva Emotional Competence (GECO) test.

5.6.1 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 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 of 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.

The analysis will be conducted in two steps using the psych R package (Revelle, 2021). 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 computed factors.

The scree test pictured in Figure 5.8, based on a factor analysis, suggests that three factors account for a fair amount of variance, with small gain achieved by adding a fourth or fifth one.

An exploratory factor analysis with three factors, using the Ordinary Least Square method of extraction with the *oblimin* rotation, result 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 5.9 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.

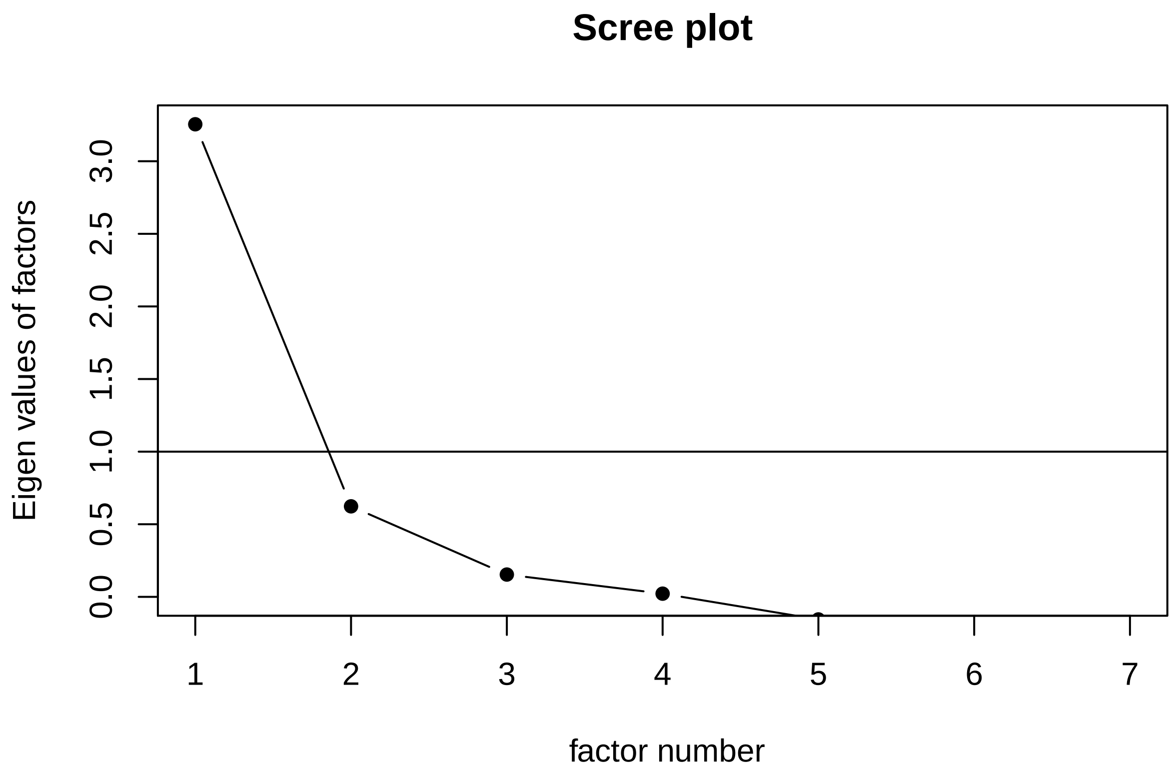


Figure 5.8: Scree plot based on all the EAU surveys administered. $N = 114$

5.6.2 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 performed 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

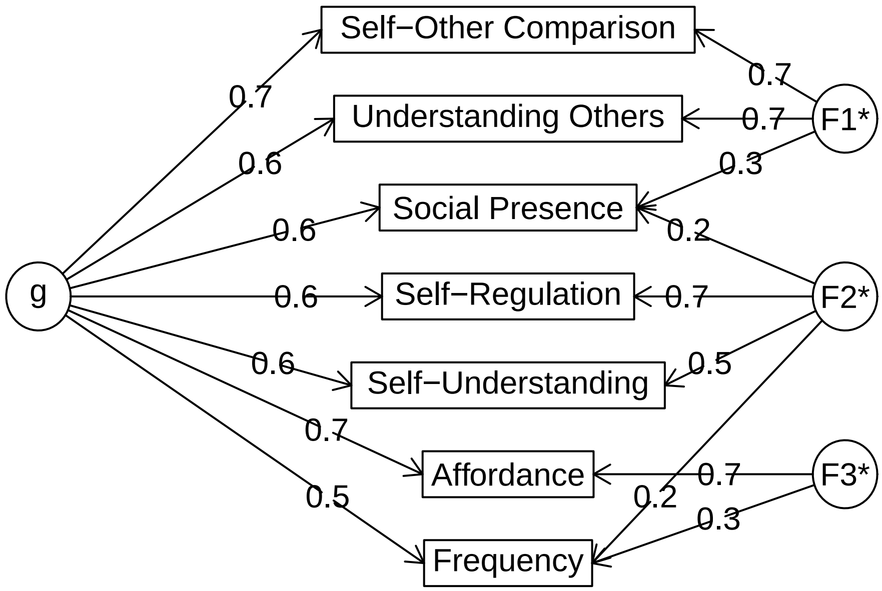


Figure 5.9: Graphical representation of the exploratory factor analysis

measure at hand to be related to emotional competence is the prospected usefulness measured in the *Expectancy* survey using the EAU scale. Even though the previous exploratory factor analysis suggests the presence of three factors in the scale, the scree test also highlights that only one factor is a reasonable compromise considering the data at hand.

These are the premises to a multiple regression that has been fit using the average score to the EAU on the *Expectancy* survey as dependent variable, and the score to the four sub-competences of the GECo test as covariates. The aim is to explore to what extent the sub-competences are related to the expected usefulness of an EAT as support to distance learning. The parameters of the multiple regression are depicted in Table 5.5, whereas the full model yielded the following results $R^2 = .60$, 90% CI [0.00, 0.75], $F(4, 6) = 2.23$, $p = .181$.

As one may expect from the small sample size at hand, there are no detectable effects in the model, with wide confidence interval around all predictors. Emotion recognition, understanding and regulation have negative estimates, whereas emotion

Table 5.5: Contrasts between the Final and the Expectancy surveys for each of the 7 dimensions of the EAU scale averaged over the two classes.

predictor	estimate	ci	statistic	p.value
Intercept	6.76	$[-0.22, 13.74]$	2.37	.056
Recognition	-4.89	$[-13.00, 3.22]$	-1.47	.191
Understanding	-0.22	$[-6.23, 5.80]$	-0.09	.933
Regulation	-0.26	$[-9.12, 8.61]$	-0.07	.945
Management	6.34	$[-0.17, 12.85]$	2.38	.055

management has a positive estimate, with a lower bound of the confidence interval that is very close to the 0 threshold. These results must obviously be taken with much more than a grain of salt, but if they were to be extended in a similar way, the fact that emotion management is the sub-competence more related to the overall expectation of the usefulness of an EAT would be an interesting result. Emotion management is in fact the subcompetence more oriented towards interpersonal regulation of emotion, that is, modifying the emotional state of others. That would imply

5.7 Conclusion

Chapter 6

Overall and Comparative Assessment of the Emotional Awareness Tool Between Learning Settings

This chapter provides an overall and comparative assessment about the use and perception of the Emotion Awareness Tool (EAT) using data collected in the two previous empirical chapters of the thesis. In particular, the chapter aims at determining (1) to what extent participants exploited the theory-driven features of the tool, including the emotional structure provided by the underlying affective space (*i.e.*, the EAT-MINT circumplex adopted in both empirical contributions); and (2) whether there are observable differences in the use and perception of the tool with respect to the prototypical distance learning settings, that is, the Asynchronous and Individual setting (Asynch./Indiv.) of the previous chapter, and the Synchronous and Collaborative setting (Synch./Collab.) of chapter 4. Naturally, the comparison between distance learning settings is not – and cannot be – aimed at generalizing the use and perception of the EAT to the specific distance learning condition (*e.g.*, assess whether a particular feeling occurs more often when learners work individually rather than in pairs). This is clearly an overarching assessment that falls way outside the scope and reach of the present contribution. Nonetheless, it is possible to take advantage of the two empirical studies to explore if the multipurpose vocation of the EAT can be sustained in two very different distance learning conditions.

In this regard, the chapter proposes four types of assessment related either to specific features of the EAT or its perception as a whole. First, the chapter explores the use of the two appraisal dimensions Valence and Control/Power that has been done through the two sliders of the Dynamic Emotion Wheel (DEW). Second, it assess the kind and proportion of subjective feelings expressed, either as part of the underlying affective space or not. Third, it evaluates the algorithm dynamically linking appraisal dimensions and subjective feelings in the adopted affective space. And last, it compares the perceived usability of the tool, which comprises the efficacy, efficiency and satisfaction in using the EAT (Brooke, 1996; Lewis & Sauro, 2018; Tullis & Albert, 2013).

6.1 Assessment of 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, 2009). As a reminder, the EATMINT circumplex adopts the Valence and Control/Power appraisal dimensions to prompt the evaluation of the eliciting event. In both empirical contributions, the Valence dimension was prompted with the question *Is the situation pleasant?*, whereas the Control/Power dimension with the question *Is the situation under your control?*. 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.

6.1.1 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 6.1 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.

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

Table 6.1: 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.	35	483	6.25 (60.76)	-4.12 (63.54)
Total	**61**	**857**	**3.55 (60.05)**	**-0.05 (58.20)**

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 divergence 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 6.1, which shows the density of the two appraisal dimensions for each learning setting. The plots highlight fairly normal 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.

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.

6.1.2 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. 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

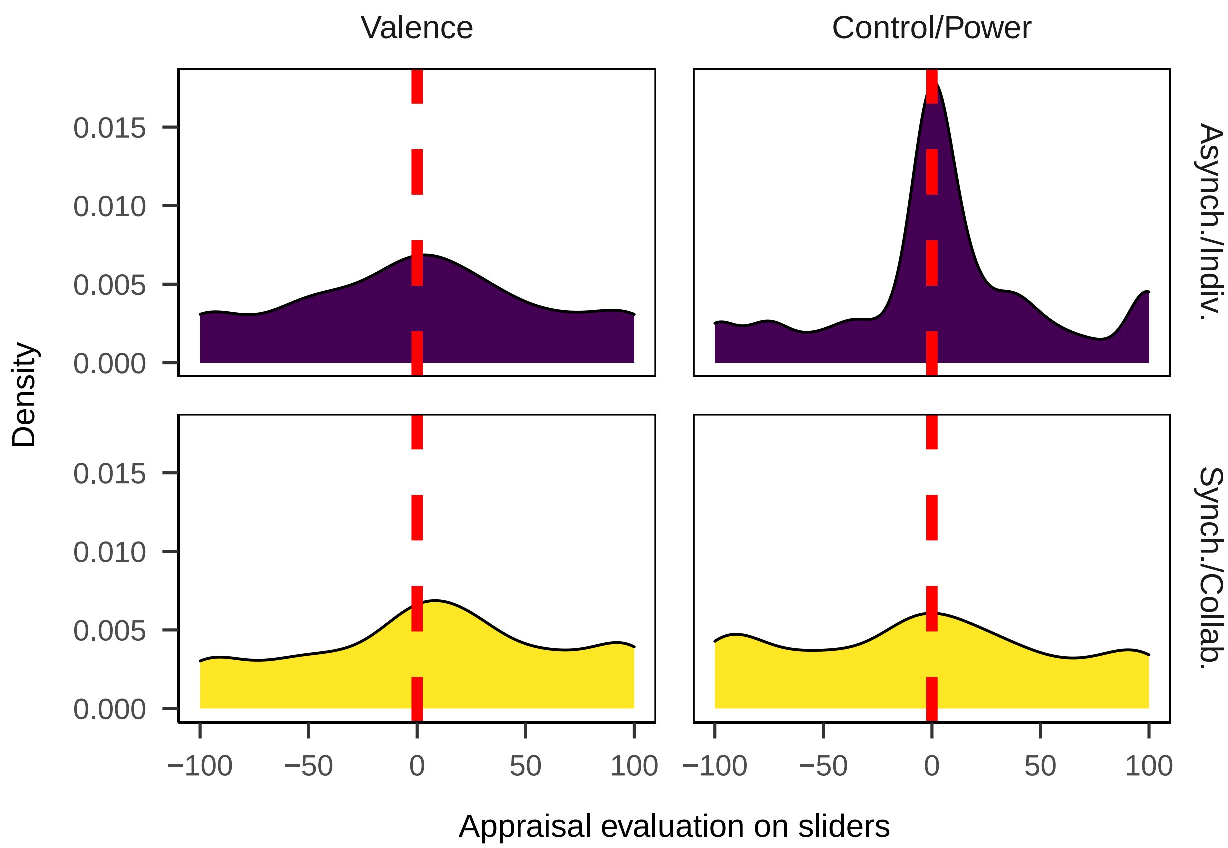


Figure 6.1: Density plots of the two appraisal dimensions' ratings for the each distance learning setting.

feeling control over it? Figure 6.2 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 6.2 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.

For both settings, the same sign combination was expressed at least one time by a

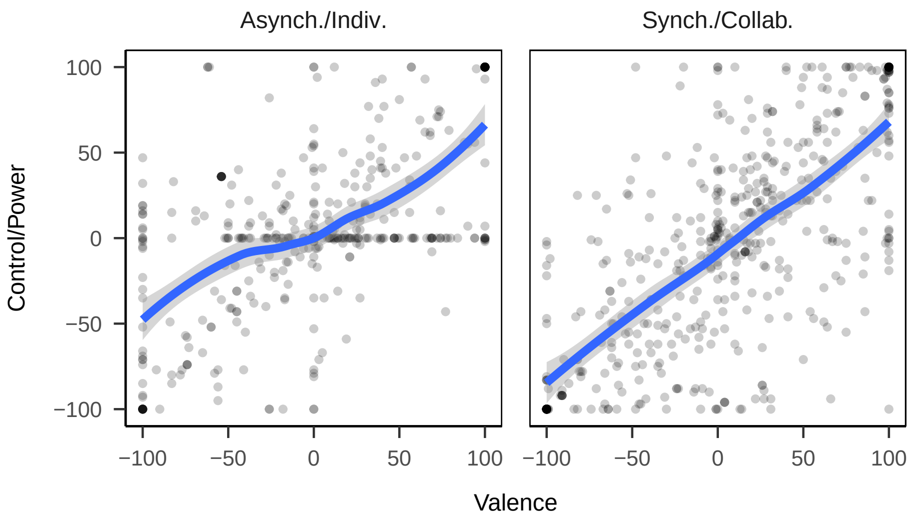


Figure 6.2: LOESS functions applied to Valence x Control/Power appraisals in the two distance learning settings.

Table 6.2: 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	34	348	0.72
Opposite sign	30	104	0.22
Either or both neutral	15	31	0.06

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.

6.1.3 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 distance learning 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.

6.2 Assessment of 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. (Grandjean, Sander, & Scherer, 2008). The assessment primarily aims at determining to what extent the feelings included in the underlying affective space – that is, the labels that appeared on the buttons or in the dropwdown menu in the expressing-displaying part of the EAT – met participants' need in terms of representation and differentiation of the conscious experience of the emotion (Barrett, Gross, Christensen, & Benvenuto, 2001; Erbas, Ceulemans, Koval, & Kuppens, 2015). As a reminder, the EATMINT circumplex proposes 20 subjective feelings, 5 for every combination of the two appraisal dimensions, which have been derived from previous studies (Fritz, 2015; Molinari, Chanel, Bétrancourt, Pun, & Bozelle, 2013). The 20 subjective feelings are Amused, Annoyed, Anxious, Attentive, Bored, Confident, Confused, Delighted, Disappointed, Disgusted, Empathetic, Envious, Frustrated, Interested, Irritated, Relaxed, Relieved, Satisfied, Stressed, and Surprised. If these 20 subjective feelings meet learners' need in best describing their conscious experience, 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 feelings of the underlying affective space are consistent with participants needs in the two distance learning settings. Table 6.3 illustrates the cumulative number of subjective feelings listed or not listed in the underlying affective space.

Data show that in both settings, participants privileged the subjective feelings 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 feelings 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 (the list of the original natural language words or idioms, in french, proposed by each setting are included in the Appendices). These results suggest that, in the Asynch./Indiv. setting, participants may need a richer *emotional vocabulary* to express their conscious emotional experience, even though the 20 subjective feelings 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 the feelings listed in the circumplex varies across distance learning settings. Table 6.4 reports the relative frequency of each one of the 20 subjective feelings in the EATMINT circumplex for both distance learning settings, as well as the absolute difference across learning settings. The use of the absolute difference highlights the fact that this comparison does not aim at determining whether participants in one learning 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 feelings with a high relative frequency in one setting and a low relative frequency in the other

Table 6.3: 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.	483	21	462	0.96
Total	**857**	**83**	**774**	**0.90**

Table 6.4: Relative frequencies of the 20 listed subjective feelings and absolute differences between distance learning settings

Feeling	Total	Asynch./Indiv.	Synch./Collab.	|Difference|
Quadrant I. Positive Valence x Positive Control/Power				
Confident	4.78	5.77	4.11	1.66
Interested	10.72	11.54	10.17	1.37
Amused	12.14	2.88	18.40	15.51
Delighted	4.78	5.13	4.55	0.58
Attentive	10.34	11.22	9.74	1.48
Quadrant II. Positive Valence x Negative Control/Power				
Satisfied	3.75	7.69	1.08	6.61
Relaxed	1.55	1.60	1.52	0.09
Surprised	4.52	0.32	7.36	7.04
Relieved	5.17	10.58	1.52	9.06
Empathetic	0.65	0.00	1.08	1.08
Quadrant III. Negative Valence x Negative Control/Power				
Confused	8.53	5.13	10.82	5.69
Anxious	3.88	4.81	3.25	1.56
Bored	8.14	8.33	8.01	0.32
Stressed	4.52	6.73	3.03	3.70
Disappointed	5.68	6.41	5.19	1.22
Quadrant IV. Negative Valence x Positive Control/Power				
Frustrated	6.98	8.33	6.06	2.27
Envious	0.13	0.00	0.22	0.22
Disgusted	0.65	0.64	0.65	0.01
Annoyed	1.42	1.92	1.08	0.84
Irritated	1.68	0.96	2.16	1.20

(*e.g.*, Amused, Relieved, Surprised or Satisfied). Second, there are feelings 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, Empathic). Results therefore corroborate the assumption that the EATMINT circumplex may adapt to different distance learning settings, even though some of its feelings are seldom chosen by participants. Whether these feelings should continue to be proposed as choices in the affective space is discussed at length in the concluding part of the contribution.

6.3 Assessment of the Link Between Appraisal Dimensions and Subjective Feelings

The two previous sections highlight that, separately, the appraisal dimensions and the subjective feelings composing the affective space seem to adapt to the two distant learning 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 model that suggests a subset of subjective feelings given a specific evaluation of the eliciting event according to *Valence* and *Control/Power* 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 (see above).

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 subjective feelings proposed in the subset of buttons on the interface, rather than having to recur to the drop-down menu or typing the word 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 with the means of *Valence* and *Control/Power* every time a given subjective feeling has been chosen by learners.

6.3.1 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 subjective feeling, 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 subjective feeling as the *right* representation, or *best* approximation, of their holistic emotional experience. 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 for each participant that has expressed at least five emotions, that is $N = 14$ in the Asynch./Indiv. setting, and $N = 33$ in the Synch./Collab. setting. Results are shown in Table 6.5.

Overall, the frequency of clicks on one of the suggested subjective feelings is $M = 0.82$ ($SD = 0.31$), which means that around 4 out of 5 emotions are expressed using one of the subjective feelings 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 $M = 0.88$, with a standard deviation of $SD = 0.28$. In both cases, thus, the frequency is well above the 0.5 threshold, which may be considered a random indicator whether the subjective feeling appeared or not in the proposed subset. 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 6.3, 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 heterogenous 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.

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 seems 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 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.

Table 6.5: Frequency of clicks on one of the suggested subjective feelings

condition	n	mean	sd
Asynch./Indiv.	14	0.70	0.36
Synch./Collab.	33	0.88	0.28
Total	**47**	**0.82**	**0.31**

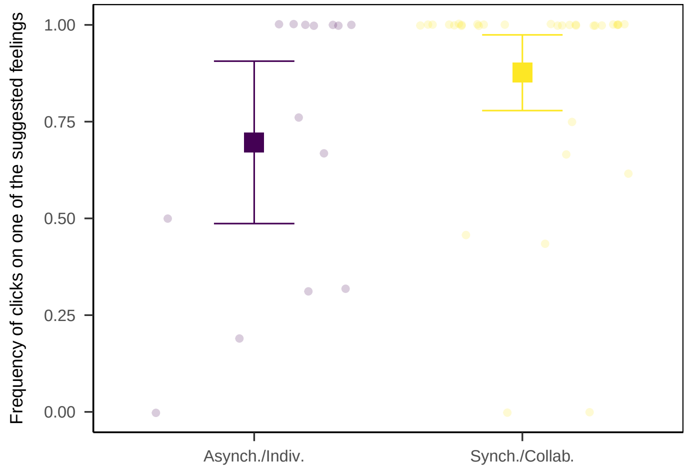


Figure 6.3: Frequency of click on one of the three buttons labeled with a subjective feeling. Bars represent 95% confidence intervals.

6.3.2 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 feeling 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 place the feeling on the circumplex. Table 6.6 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 6.6: 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	36	15.86 (28.14)	41.67 (35.52)	20.84	9	11.84
Interested	83	45.67 (35.58)	47.42 (36.83)	43.92	27	16.92
Amused	94	49.64 (41.62)	38.50 (51.07)	52.20	45	7.20
Delighted	37	72.19 (33.18)	65.03 (38.30)	47.99	63	15.01
Attentive	80	32.16 (37.32)	18.75 (31.03)	59.76	81	21.24
Quadrant II. Positive Valence x Negative Control/Power						
Satisfied	29	59.34 (38.00)	30.86 (39.72)	62.52	99	36.48
Relaxed	12	56.25 (39.23)	11.00 (42.15)	78.94	117	38.06
Surprised	35	33.29 (33.90)	-27.23 (40.30)	129.28	135	5.72
Relieved	39	47.62 (43.23)	7.62 (52.14)	80.91	153	72.09
Empathetic	5	1.80 (7.40)	-57.60 (39.95)	178.21	171	7.21
Quadrant III. Negative Valence x Negative Control/Power						
Confused	66	-18.18 (37.93)	-49.89 (34.43)	200.02	189	11.02
Anxious	30	-42.73 (33.28)	-42.30 (41.19)	225.29	207	18.29
Bored	63	-55.95 (36.53)	-55.52 (40.50)	225.22	225	0.22
Stressed	35	-58.14 (35.87)	-36.49 (42.94)	237.89	243	5.11
Disappointed	44	-52.73 (40.22)	-19.82 (39.69)	249.40	261	11.60
Quadrant IV. Negative Valence x Positive Control/Power						
Frustrated	54	-45.33 (38.06)	-26.91 (47.16)	239.31	279	39.69
Envious	1	-52.00 (NA)	25.00 (NA)	295.68	297	1.32
Disgusted	5	-75.00 (26.80)	-44.40 (69.00)	239.37	315	75.63
Annoyed	11	-72.45 (38.59)	-59.91 (68.67)	230.41	333	102.59
Irritated	13	-25.54 (23.25)	43.31 (53.48)	329.47	351	21.53

The results highlight a wide range of differences between the expected and the observed disposition of each feeling, 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 feelings is computed using only a few observations, such 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, with respect to the use of the two appraisal dimensions. In fact, comparing the graphical representations of the theoretical/expected circumplex in Figure 6.4 with the empirical/observed in Figure 6.5 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 feelings of the orthogonal combination of appraisals moved either in the *upper* or *lower* 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 6.6 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 algorithm. As Figure 6.6 shows, the overall tendency of co-variation in the two appraisal dimensions permains. Nevertheless, in the Synch./Collab. setting, there is more consistency in the bottom-right quadrant, the one charactherized by Positive Valence x Negative Control. 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.

6.3.3 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 these conditions.

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

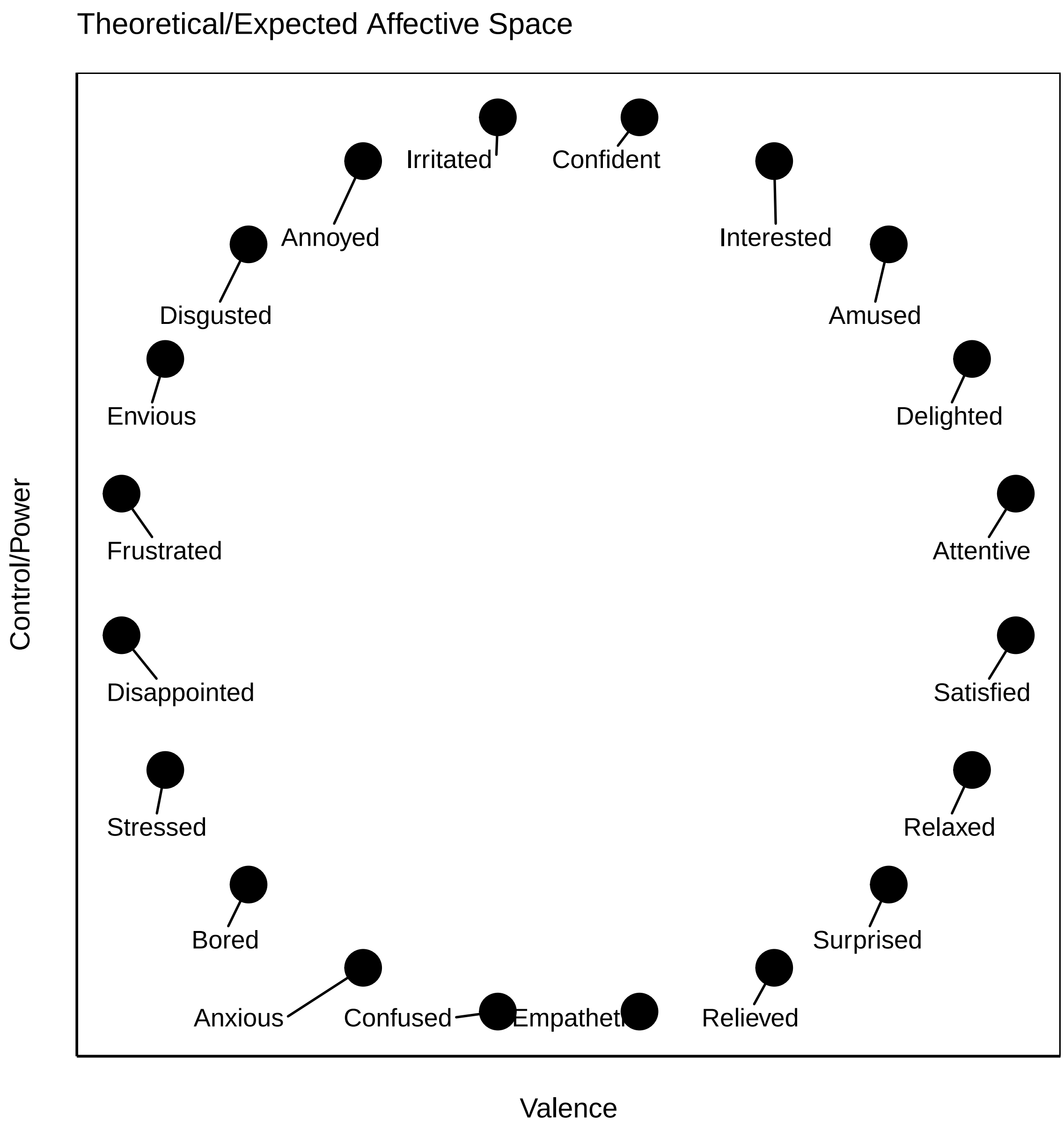


Figure 6.4: The theoretical/expected disposition of the EATMINT circumplex.

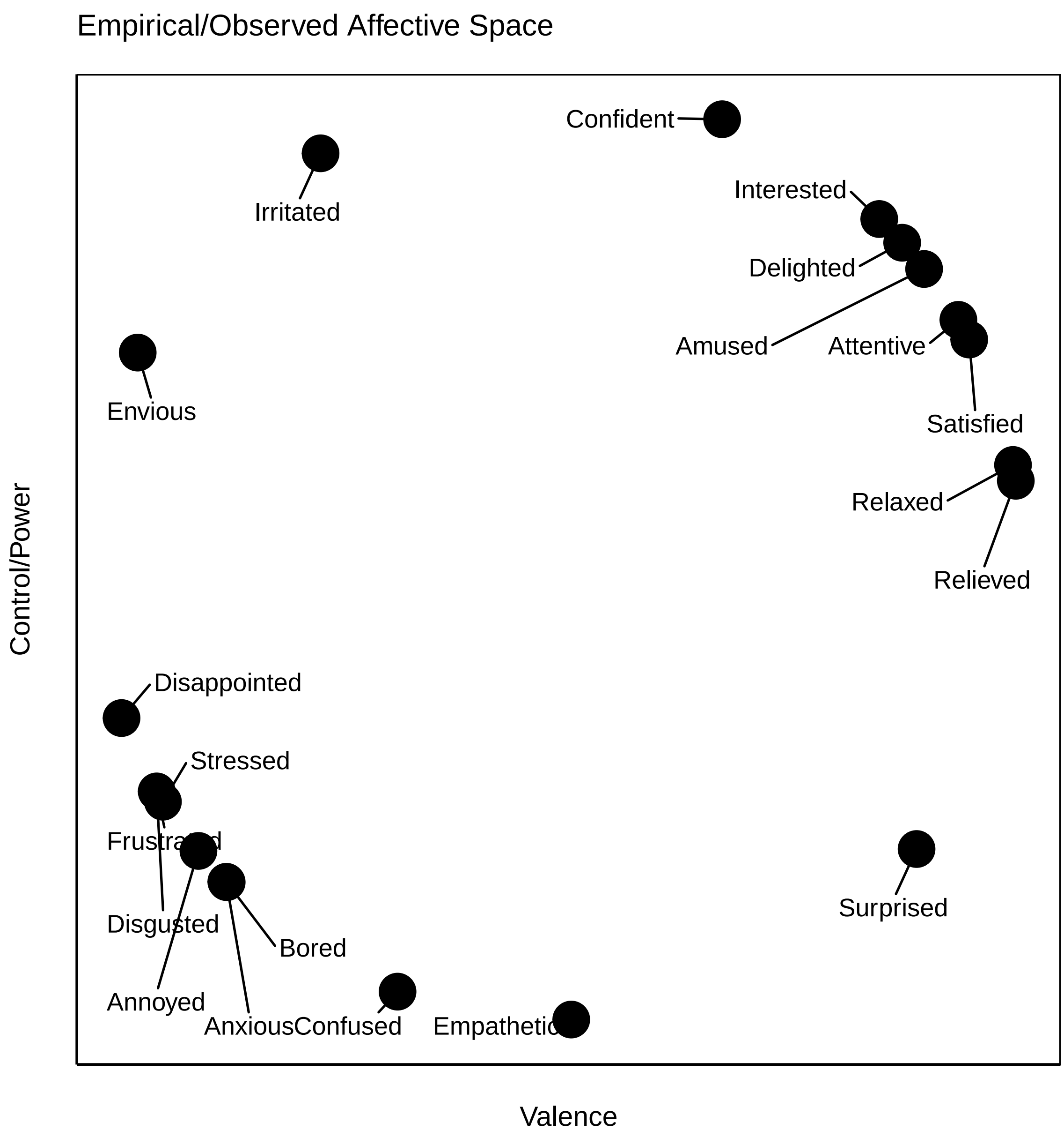


Figure 6.5: The empirical/observed disposition of the feelings according to the actual rate of participants.

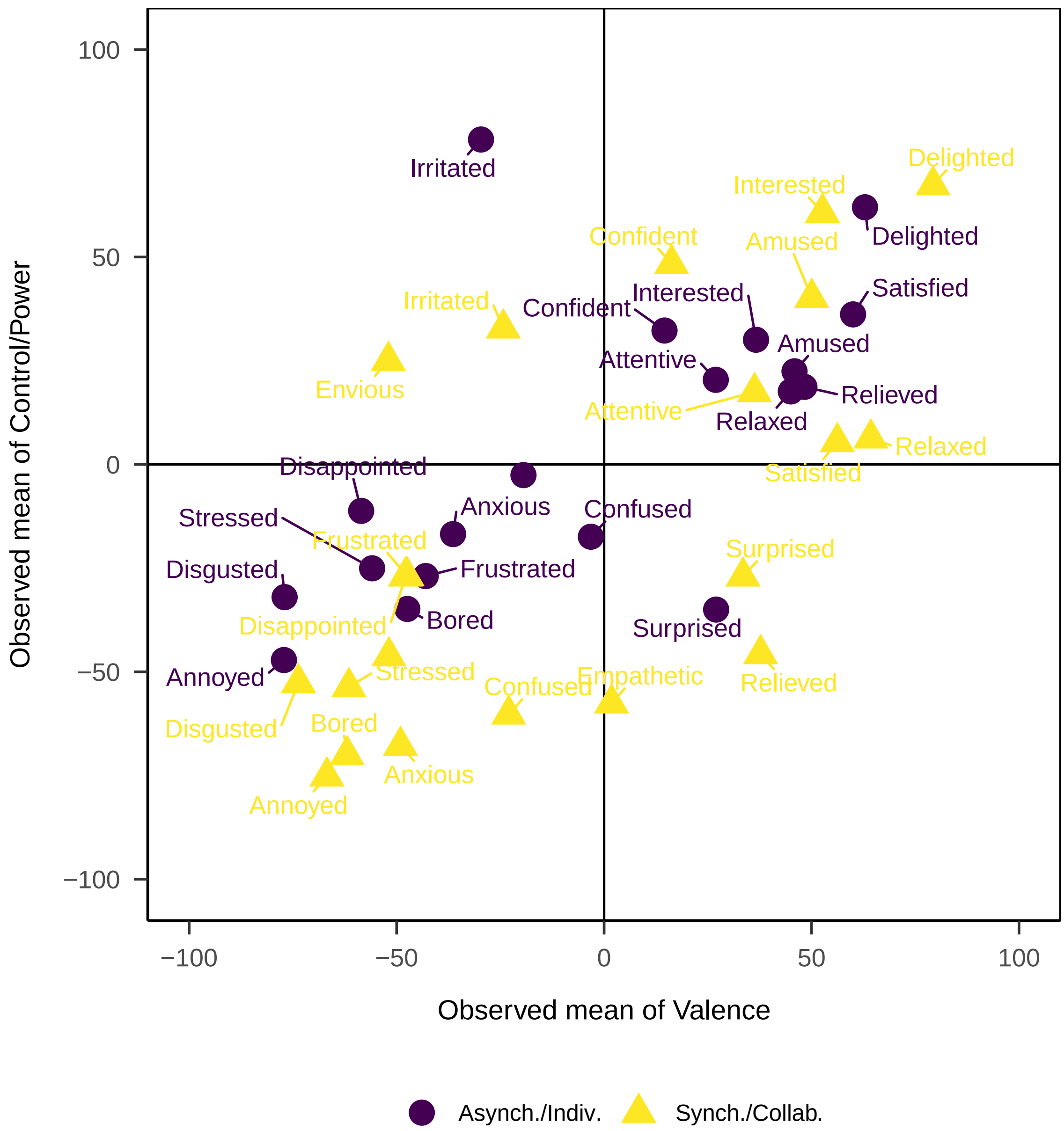


Figure 6.6: Comparing the empirical disposition of the two learning settings.

phenomenon already observed in the usability test of the DEW (Fritz, 2015). The problem, though, does not seem to be unique to the tool. On the contrary, Scherer and Fontaine (2018) encountered a similar difficulty with the use of the GRID instrument (Fontaine, Scherer, & Soriano, 2013a), which, as a reminder, consists in a questionnaire comprising 142 features related to the different components of an emotion (*i.e.*, appraisal, bodily symptoms, expression, action tendencies, and subjective experience). Scherer and Fontaine (2018), in an attept to investigate whether the dimensions of the feeling component could be predicted by the other components, among which the appraisal, failed to make emerge in the results the importance of the Control/Power dimension. In this regard, the authors note (*ibid.*, p. 9) :

> 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).

As advocated by the authors, this problem requires future studies to find better solution. In this regard, some options pertaining to the DEW will be provided in the general discussion of the thesis.

6.4 Assessment of 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 Asynch./Indiv. empirical contribution, but not in the Synch./Collab. The SUS was nevertheless administered in the usability test in Fritz (2015), which used the same configuration and task of the Synch./Collab. 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. setting of the present contribution ($N = 26$).

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, Kortum, & Miller, 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 $M = 68$, but 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 6.7 shows the SUS score for the two learning settings, as well as the weighted overall mean. With an overall score of $M = 73.10$ ($SD = 12.74$), 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 6.7: Score to the System Usability Scale (SUS, Brooke, 1996)

Condition	N	M	SD
Asynch./Indiv.	26	72.82	11.89
Synch./Collab.	16	73.54	14.43
Total	**42**	**73.10**	**12.74**

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 6.7 depicts the score of each of the SUS items across learning 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.

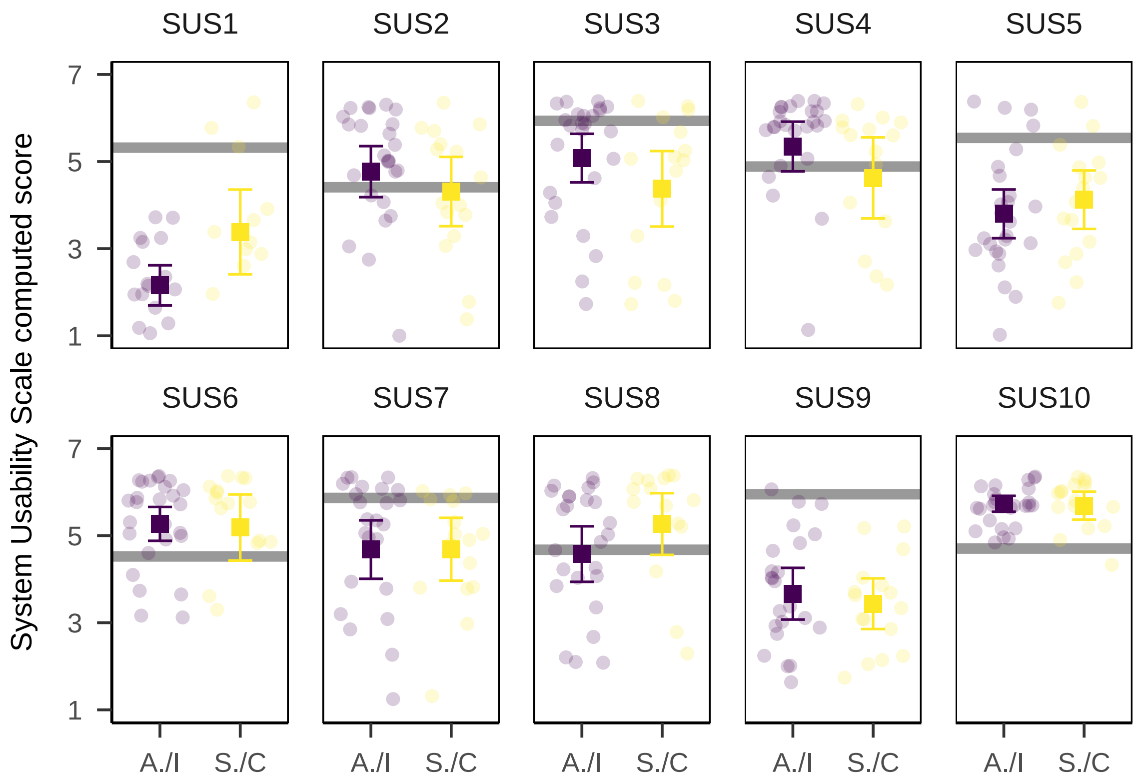


Figure 6.7: SUS scores on the single items, with the horizontal lines representing Lewis and Sauro (2018) benchmarks to reach the target score of 80, transformed to a 1-to-7 scale. Both items and benchmarks have already been reversed for even items. Bars represent 95% confidence intervals.

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 distance learning 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.

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.

The implications of the usability evaluation of the EAT will be integrated in the general discussion of the thesis, since they provide useful information that can be linked to broader aspects about emotional awareness and the tool more specifically. A preliminary synthesis nevertheless 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.

6.5 Conclusion

This chapter proposed an overall assessment of the use and perception of the EAT, as well as a comparison between the use in an asynchronous and individual distance learning environment, versus a synchronous and collaborative one. The primary purpose of the chapter 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 distance learning settings. For instance, the 20 subjective feelings proposed in the circumplexes 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 accuracy, consistent across learning settings. These indicators suggest that the EAT possesses a sort of *intrinsic* value, which may be generalized 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 suggest 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 in the first test of the tool (Fritz, 2015), such as the lack of orthogonality of the two appraisal dimensions Valence and Control/Power, or issues with the prospected frequency of use and the confidence in the system. The issues, though, do not seem limited to technical elements of the EAT, but rather extend to more fundamental aspects of emotional awareness or the difficulty in determining an emotional structure that can be leveraged by users. As a consequence, they need a thorough and integrated perspective, which is provided next in the general discussion of the thesis.

Chapter 7

General Discussion

- Things should be simple, but not simpler
- Can an EAT really be multi-purpose
- Tenacity vs. obstination (emotional structure vs. use)
- Emotional differentiation vs. frequency
- Research vs. practice

Conclusion

Appendix A

Study 1 - Synch./Coll.

A.1 Expressing Emotions

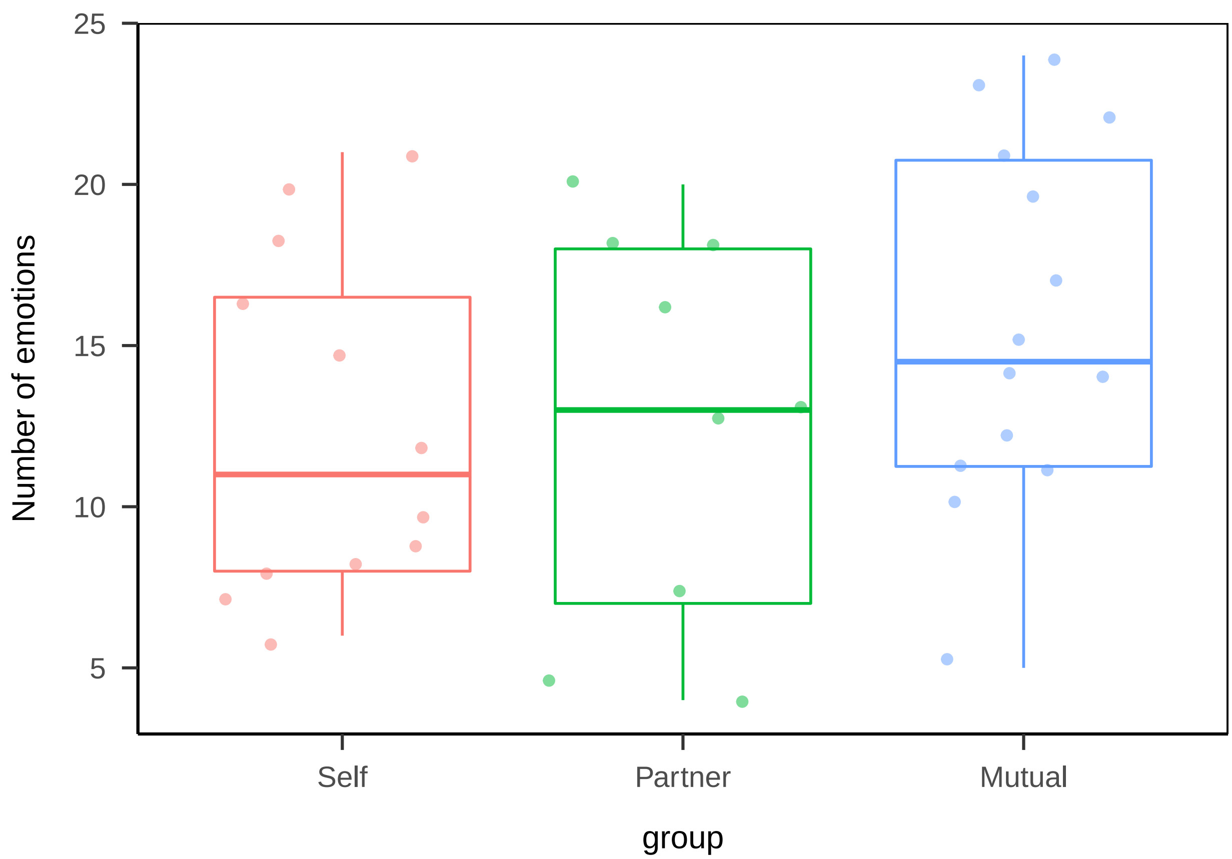


Figure A.1: Repartition of expressed emotions per condition

Table A.1: Analysis of Variance Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	79.39	39.69	1.247	0.3008
Residuals	32	1018	31.82	NA	NA

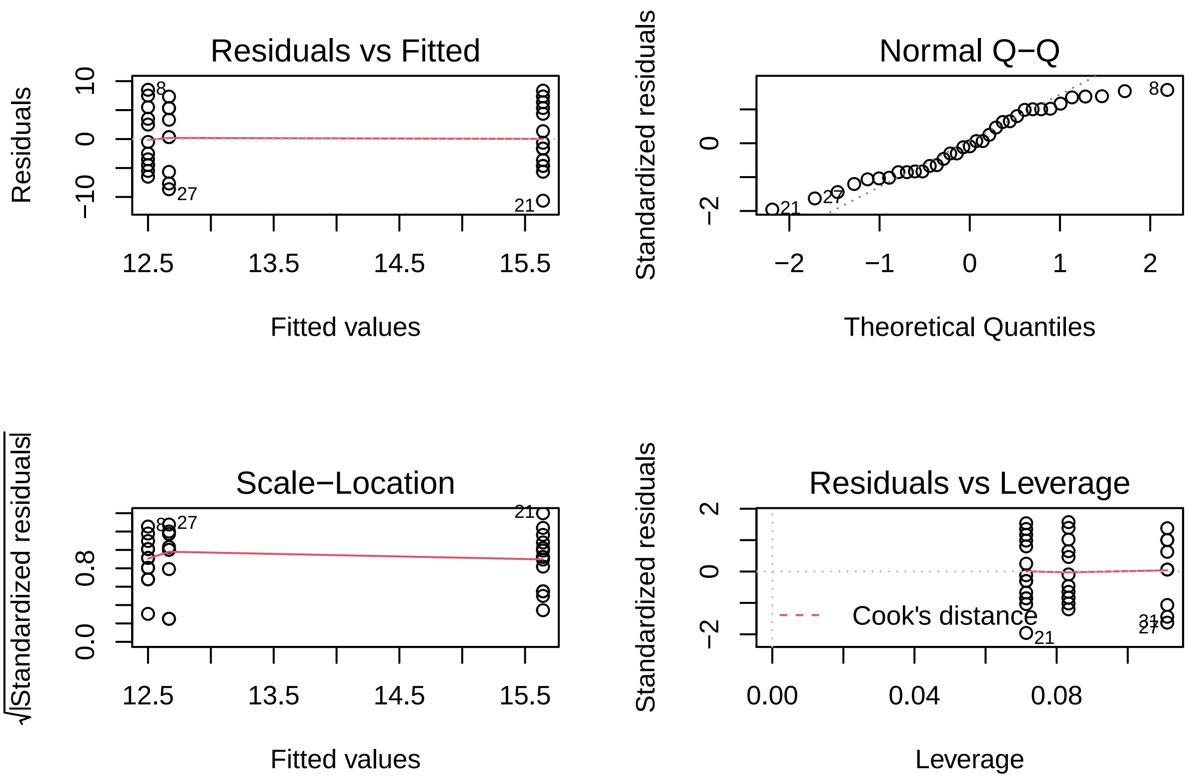


Figure A.2: Assumptions of the one-way ANOVA for expressing emotions

Table A.2: Levene's Test for Homogeneity of Variance (center = center)

	Df	F value	Pr(>F)
group	2	0.02045	0.9798
	32	NA	NA

A.2 Processing Emotional Information

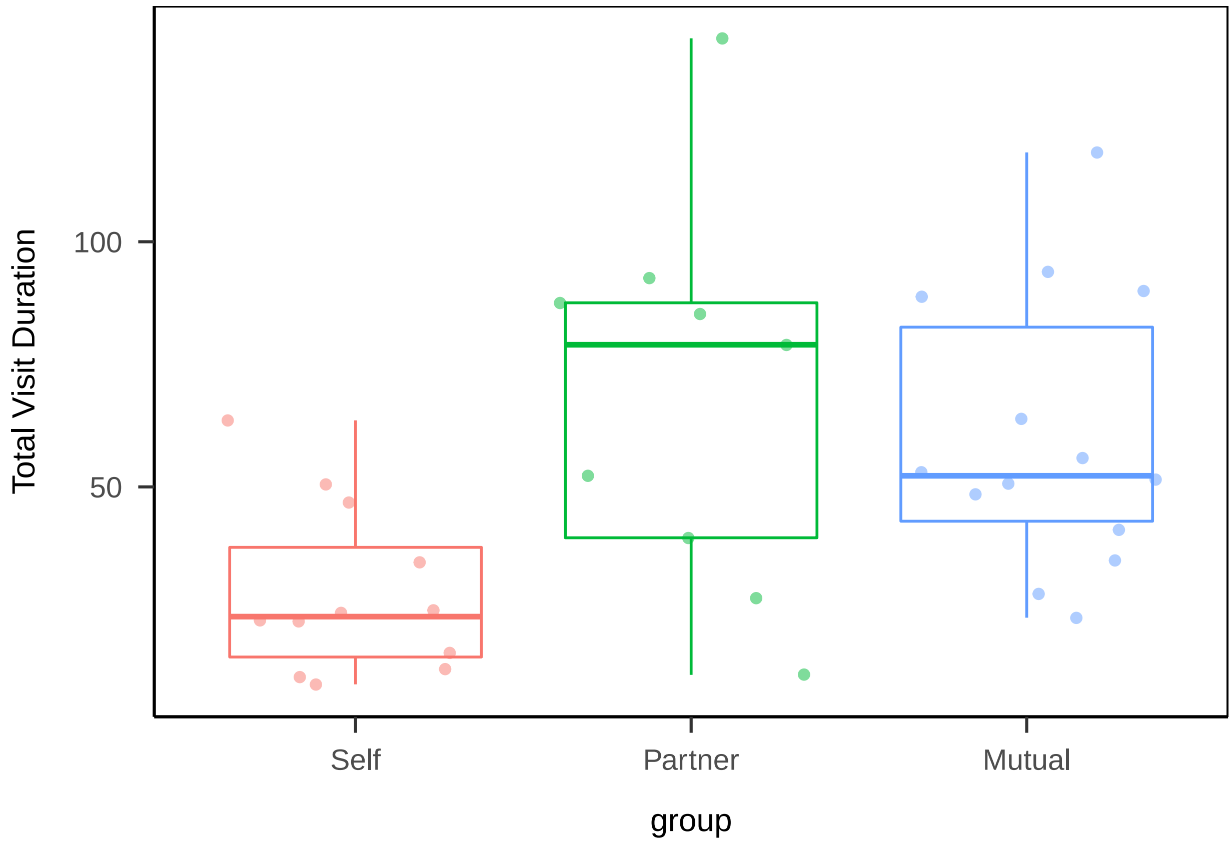


Figure A.3: Repartition of total visit duration per condition

Table A.3: Analysis of Variance Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	10097	5048	6.238	0.005157
Residuals	32	25898	809.3	NA	NA

Table A.4: Levene's Test for Homogeneity of Variance (center = center)

	Df	F value	Pr(>F)
group	2	2.298	0.1168
	32	NA	NA

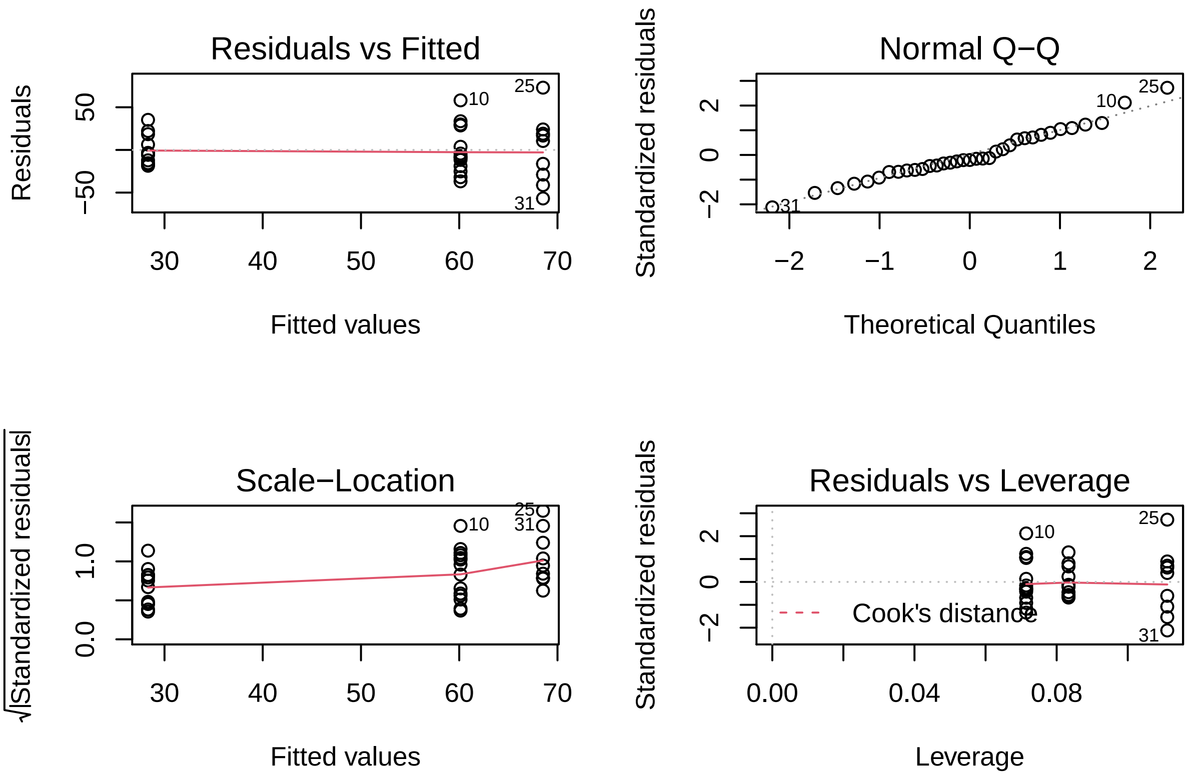


Figure A.4: Assumptions of the one-way ANOVA for processing emotional information

A.3 Seeking Emotional Information

Table A.5: Analysis of Variance Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
group	2	10097	5048	6.238	0.005157
Residuals	32	25898	809.3	NA	NA

Table A.6: Levene's Test for Homogeneity of Variance (center = center)

	Df	F value	Pr(>F)
group	2	1.767	0.1871
	32	NA	NA

A.4 Transitions Between AOI

Table A.7: Summary of the number of transitions aggregated per participant, stratified by experimental condition and path.

group	n	mean	sd	ci.lower	ci.upper
Expressing to Perceiving					
Self	12	21.33	8.38	16.01	26.66
Partner	8	22.25	8.60	15.06	29.44
Mutual	14	39.43	15.21	30.64	48.21
Perceiving to Expressing					
Self	12	18.42	9.04	12.67	24.16
Partner	8	22.75	8.65	15.52	29.98
Mutual	14	35.79	14.15	27.62	43.95
Expressing to Task					
Self	12	39.92	9.12	34.12	45.71
Partner	8	29.62	8.48	22.53	36.72
Mutual	14	32.79	11.56	26.11	39.46
Task to Expressing					
Self	12	42.75	11.05	35.73	49.77
Partner	8	29.62	7.87	23.04	36.21
Mutual	14	36.29	14.03	28.19	44.39
Perceiving to Task					
Self	12	17.17	5.78	13.49	20.84
Partner	8	24.25	10.15	15.76	32.74
Mutual	14	34.36	10.96	28.03	40.68
Task to Perceiving					
Self	12	14.42	7.91	9.39	19.44
Partner	8	24.88	11.29	15.43	34.32
Mutual	14	30.14	9.36	24.74	35.55

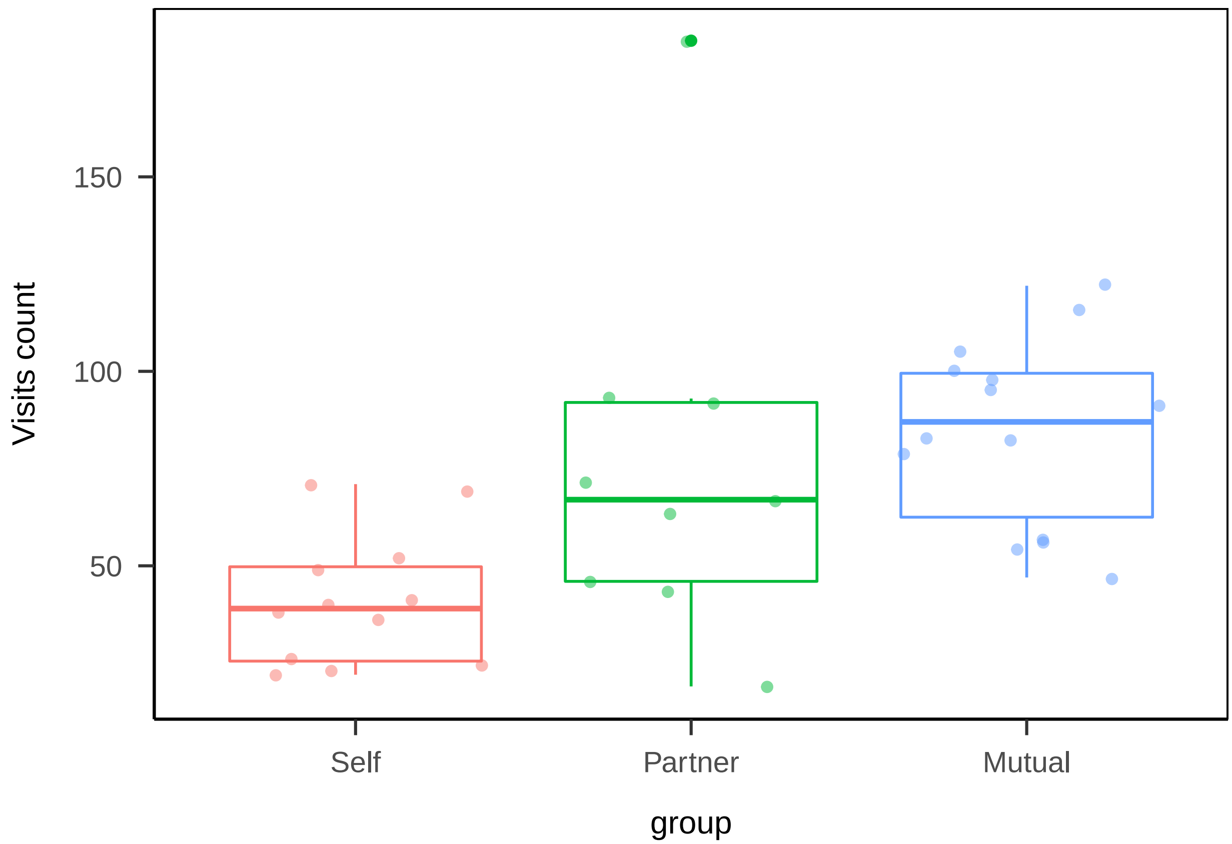


Figure A.5: Repartition of visit count per condition

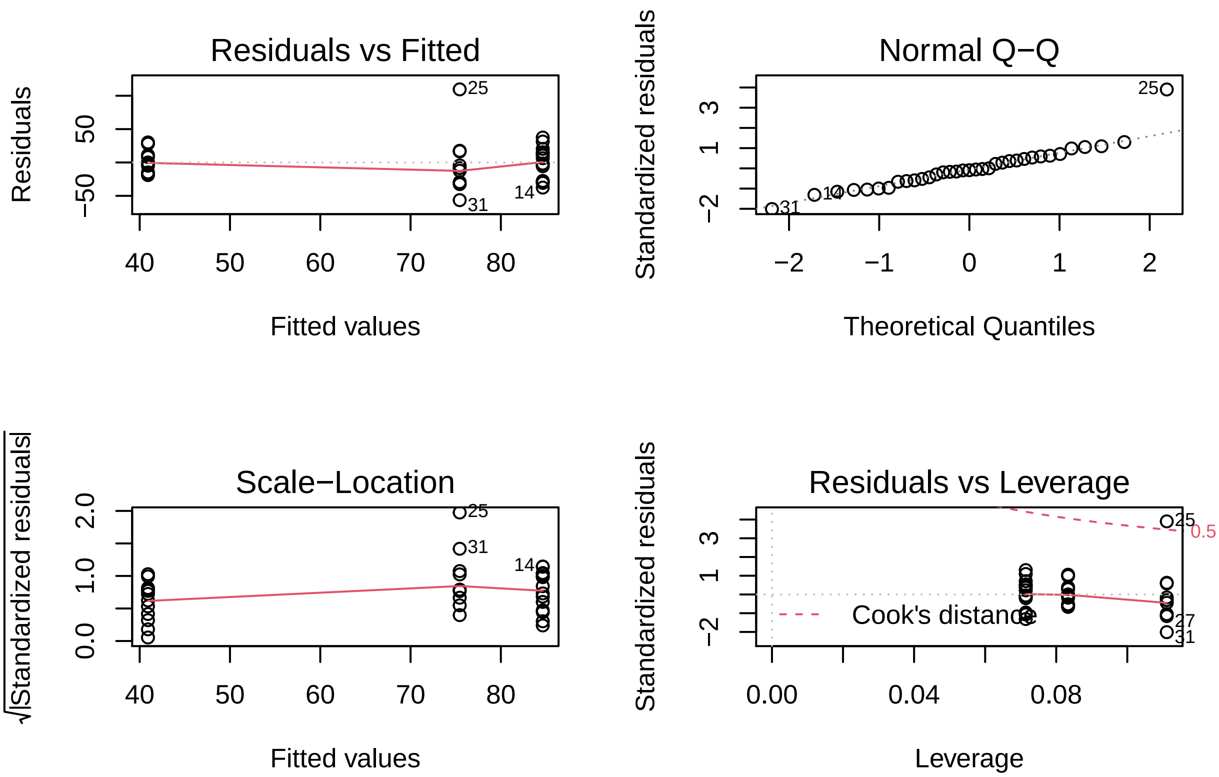


Figure A.6: Assumptions of the one-way ANOVA for seeking emotional information

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[1]

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[2]

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology*, *84*(5), 1054–1068. http://doi.org/10.1037/0022-3514.84.5.1054[3]

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[4]

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.

[1] `https://doi.org/10.1177/1754073918765662`
[2] `https://doi.org/10.1016/j.jesp.2017.09.004`
[3] `https://doi.org/10.1037/0022-3514.84.5.1054`
[4] `https://doi.org/10.1080/23743603.2019.1684822`

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[5]

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[6]

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& apos;Mello, S. K., Rodrigo, M. 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[7]

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[8]

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[9]

Barrett, L. F. (2018). *How emotions are made: The secret life of the brain* (Paperback edition). London: PAN Books.

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[10]

[5] `https://doi.org/10.1109/INCoS.2015.78`
[6] `https://doi.org/10.3389/fpsyg.2020.01160`
[7] `https://doi.org/10.1016/j.ijhcs.2009.12.003`
[8] `https://doi.org/10.1006`
[9] `https://doi.org/66.39.39.113`
[10] `https://doi.org/10.1080/02699930143000239`

Barsade, S. (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[11]

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018, May). Parsimonious Mixed Models. Retrieved from http://arxiv.org/abs/1506.04967

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014, June). Fitting Linear Mixed-Effects Models using Lme4. Retrieved from http://arxiv.org/abs/1406.5823

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[12]

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[13]

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[14]

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[15]

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.

[11] https://doi.org/10.1146/annurev-orgpsych-032414-111316
[12] https://doi.org/10.1111/cgf.13079
[13] https://doi.org/10.1177/0146167220936054
[14] https://doi.org/10.1016/j.chb.2010.07.014
[15] https://doi.org/10.1016/j.ijhcs.2006.11.016

http://doi.org/10.1080/00461520.2019.1614447[16]

Brackett, M. A., Rivers, S. E., Reyes, M. R., & Salovey, P. (2012). 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[17]

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[18]

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., 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[19]

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[20]

Butler, E. A., Gomes De Mesquita, B., & Feldman Barrett, L. (2017). Emotions are temporal interpersonal systems Models of temporal interpersonal emotion systems (TIES). *Current Opinion in Psychology*, *17*, 129–134. http://doi.org/10.1016/j.copsyc.2017.07.005[21]

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.

[16] `https://doi.org/10.1080/00461520.2019.1614447`
[17] `https://doi.org/10.1016/j.lindif.2010.10.002`
[18] `https://doi.org/10.1016/j.ijhcs.2013.02.003`
[19] `https://doi.org/10.4414/smw.2013.13786`
[20] `https://doi.org/10.1016/j.chb.2010.07.012`
[21] `https://doi.org/10.1016/j.copsyc.2017.07.005`

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[22]

Calvo, R., D'Mello, S., Gratch, J., Kappas, A., D'Mello, S., & Graesser, A. C. (2015). Feeling, Thinking, and Computing with Affect-Aware Learning Technologies. In R. 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[23]

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[24]

Carswell, L., Thomas, P., Petre, M., Price, B., & Richards, M. (2000). Distance education via the Internet: The student experience. *British Journal of Educational Technology*, *31*(1), 29–46. http://doi.org/10.1111/1467-8535.00133[25]

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[26]

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 / Chris Chambers*. 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[27]

Chen, Y., Ma, X., Cerezo, A., & Pu, P. (2014). Empatheticons: Designing Emotion Awareness Tools for Group Recommenders. In *Proceedings of*

[22] `https://doi.org/10.1109/MIS.2012.110`
[23] `https://doi.org/10.1093/oxfordhb/9780199942237.013.032`
[24] `https://doi.org/10.1037/1076-898X.4.2.75`
[25] `https://doi.org/10.1111/1467-8535.00133`
[26] `https://doi.org/10.1016/j.jvlc.2015.10.001`
[27] `https://doi.org/10.1007/978-3-319-12562-6`

the XV International Conference on Human Computer Interaction - Interacción '14 (pp. 1–8). Puerto de la Cruz, Tenerife, Spain: ACM Press. http://doi.org/10.1145/2662253.2662269[28]

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[29]

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[30]

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[31]

Conrad, D. L. (2002). Engagement, excitement, anxiety, and fear: Learners' experiences of starting an Online course. *American Journal of Distance Education*, *16*(4), 205–226. http://doi.org/10.1207/S15389286AJDE1604_2[32]

Craig, S. D., D'Mello, S., Witherspoon, A., & Graesser, A. (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[33]

Crüwell, S., van Doorn, J., Etz, A., Makel, M. C., Moshontz, H., Niebaum, J. C., ... Schulte-Mecklenbeck, M. (2019). Seven Easy Steps to Open Science: An Annotated Reading List. *Zeitschrift für Psychologie*, *227*(4), 237–248. http://doi.org/10.1027/2151-2604/a000387[34]

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[35]

[28] https://doi.org/10.1145/2662253.2662269
[29] https://doi.org/10.1111/j.1754-9434.2010.01222.x
[30] https://doi.org/10.1016/j.obhdp.2011.06.002
[31] https://doi.org/10.1037/10096-006
[32] https://doi.org/10.1207/S15389286AJDE1604_2
[33] https://doi.org/10.1080/02699930701516759
[34] https://doi.org/10.1027/2151-2604/a000387
[35] https://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[36]

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. http://doi.org/10.1016/j.learninstruc.2011.10.00

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. http://doi.org/10.1016/j.learninstruc.2012.05.003[38]

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.

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[39]

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[40]

Desmet, P. (2003). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In *Funology* (pp. 111–123). Springer.

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[41]

[36] `https://doi.org/10.1037/a0032674`
[37] `https://doi.org/10.1016/j.learninstruc.2011.10.001`
[38] `https://doi.org/10.1016/j.learninstruc.2012.05.003`
[39] `https://doi.org/10.1016/j.chb.2010.05.018`
[40] `https://doi.org/10.1016/j.chb.2007.04.004`
[41] `https://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[42]

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[43]

Dixon, T. (2012). Educating the emotions from Gradgrind to Goleman. *Research Papers in Education*, *27*(4), 481–495. http://doi.org/10.1080/02671522.2012.690240[44]

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[45]

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.

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[46]

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[47]

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[48]

[42] https://doi.org/10.1007/978-1-4020-9827-7_1
[43] https://doi.org/10.1007/s11412-016-9235-5
[44] https://doi.org/10.1080/02671522.2012.690240
[45] https://doi.org/10.1145/143457.143468
[46] https://doi.org/10.1016/j.chb.2009.04.004
[47] https://doi.org/10.1037/emo0000039
[48] https://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[49]

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[50]

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[51]

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[52]

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[53]

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (2nd ed). Hoboken, N.J: Wiley.

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

[49] `https://doi.org/10.1504/IJCEELL.2014.060154`
[50] `https://doi.org/10.1177/1745691612459059`
[51] `https://doi.org/10.1177/2515245918770407`
[52] `https://doi.org/10.1177/1754073910361980`
[53] `https://doi.org/10.1080/00461520.2012.748005`

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[54]

Fontaine, J. R. J., Scherer, K. R., & Soriano, C. (Eds.). (2013a). *Components of Emotional Meaning*. Oxford, UK: Oxford University Press.

Fontaine, J. R. J., Scherer, K. R., & Soriano, C. (2013b). The why, the what, and the how of the GRID instrument. In J. R. J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp. 83–97). Oxford, UK: Oxford University Press.

Fox, A. S. (Ed.). (2018). *The nature of emotion: Fundamental questions* (Second edition). New York, NY: Oxford University Press.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge 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. (2019). An Overview of the Replication Crisis in Psychology and a Project that May Help.

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.

[54] `https://doi.org/10.1111/j.1467-9280.2007.02024.x`

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.

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[55]

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[56]

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[57]

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[58]

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences, 281–291. http://doi.org/10.1017.S0048577201393198[59]

Gross, J. J. (Ed.). (2014). *Handbook of emotion regulation* (2. ed). New York, NY: Guilford Press.

[55] `https://doi.org/10.1080/02699931.2017.1287668`
[56] `https://doi.org/10.1024/1421-0185/a000180`
[57] `https://doi.org/10.1177/1948550617720269`
[58] `https://doi.org/10.1007/s11257-017-9188-z`
[59] `https://doi.org/10.1017.S0048577201393198`

Gross, J. J. (2015). Emotion Regulation: Current Status and Future Prospects. *Psychological Inquiry*, (26), 1–26.

Grudin, J. (1994). Computer-Supported Cooperative Work: History and Focus. *IEEE Computer*, *27*, 19–26. http://doi.org/10.1109/2.291294[60]

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[61]

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[62]

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[63]

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[64]

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[65]

Harley, Jason Matthew. (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, Jason M., Lajoie, S. P., Frasson, C., & Hall, N. C. (2017). Developing Emotion-Aware, Advanced Learning Technologies: A Taxonomy of Approaches

[60] `https://doi.org/10.1109/2.291294`
[61] `https://doi.org/10.1080/08923649709526970`
[62] `https://doi.org/10.1023/A:1021271517844`
[63] `https://doi.org/10.1145/257089.257286`
[64] `https://doi.org/10.1080/02699930802613828`
[65] `https://doi.org/10.1080/02699931.2013.791615`

and Features. *International Journal of Artificial Intelligence in Education*, *27*(2), 268–297. http://doi.org/10.1007/s40593-016-0126-8[66]

Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. (2016). The Discrete Emotions Questionnaire: A New Tool for Measuring State Self-Reported Emotions. *PLOS ONE*, *11*(8), e0159915. http://doi.org/10.1371/journal.pone.0159915[67]

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional Contagion. *Current Directions in Psychological Science*, *2*(3), 96–99.

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[68]

Hodzic, S., Scharfen, J., Ripoll, P., Holling, H., & Zenasni, F. (2018). How Efficient Are Emotional Intelligence Trainings: A Meta-Analysis. *Emotion Review*, *10*(2), 138–148. http://doi.org/10.1177/1754073917708613[69]

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[70]

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[71]

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.

Isomursu, M., Tähti, M., Väinämö, S., & Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human Computer Studies*, *65*, 404–418. http://doi.org/10.1016/j.ijhcs.2006.11.007[72]

[66] `https://doi.org/10.1007/s40593-016-0126-8`
[67] `https://doi.org/10.1371/journal.pone.0159915`
[68] `https://doi.org/10.1111/j.1756-8765.2011.01150.x`
[69] `https://doi.org/10.1177/1754073917708613`
[70] `https://doi.org/10.3389/fpsyg.2012.00137`
[71] `https://doi.org/10.1037/emo0000649`
[72] `https://doi.org/10.1016/j.ijhcs.2006.11.007`

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[73]

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[74]

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[75]

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[76]

Järvenoja, H., & Järvelä, S. (2009). Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges? *The British Journal of Educational Psychology*, *79*, 463–81. http://doi.org/10.1348/000709909X402811[77]

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[78]

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[79]

[73] `https://doi.org/10.3406/rfp.1993.1305`
[74] `https://doi.org/10.1080/00461520.2012.749153`
[75] `https://doi.org/10.1016/j.chb.2010.06.002`
[76] `https://doi.org/10.1007/s11412-016-9238-2`
[77] `https://doi.org/10.1348/000709909X402811`
[78] `https://doi.org/10.1080/01443410.2012.742334`
[79] `https://doi.org/10.1177/0956797611430953`

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[80]

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[81]

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196–217. http://doi.org/10.1207/s15327957pspr0203_4[82]

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[83]

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[84]

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[85]

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[86]

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[87]

Lavoué, E., Kazemitabar, M., Doleck, T., Lajoie, S. P., Carrillo, R., & Molinari, G. (2019). Towards emotion awareness tools to support emotion and appraisal

[80] `https://doi.org/10.1080/026999399379168`
[81] `https://doi.org/10.1146/annurev-psych-010419-051123`
[82] `https://doi.org/10.1207/s15327957pspr0203_4`
[83] `https://doi.org/10.1111/jcal.12084`
[84] `https://doi.org/10.1016/S0747-5632(02)00057-2`
[85] `https://doi.org/10.1080/00461520.2012.750225`
[86] `https://doi.org/10.18637/jss.v082.i13`
[87] `https://doi.org/10.1177/1948550617697177`

regulation in academic contexts. *Educational Technology Research and Development*. http://doi.org/10.1007/s11423-019-09688-x[88]

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[89]

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[90]

Lehman, B., D'Mello, S., & Graesser, A. (2012). Confusion and complex learning during interactions with computer learning environments. *The Internet and Higher Education*, *15*(3), 184–194. http://doi.org/10.1016/j.iheduc.2012.01.002[91]

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[92]

Levenson, R. W. (1999). The Intrapersonal Functions of Emotion. *Cognition & Emotion*, *13*(5), 481–504. http://doi.org/10.1080/026999399379159[93]

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[94]

Lewis, J. R., & Sauro, J. (2018). Item Benchmarks for the System Usability Scale. *Journal of Usability Studies*, *13*(3), 158–167.

Lieberman, Matthew D. (2019). Affect labeling in the age of social media. *Nature Human Behaviour*, *3*(1), 20–21. http://doi.org/10.1038/s41562-018-0487-0[95]

[88] https://doi.org/10.1007/s11423-019-09688-x
[89] https://doi.org/10.1007/s11423-019-09688-x
[90] https://doi.org/10.1007/s11218-016-9359-5
[91] https://doi.org/10.1016/j.iheduc.2012.01.002
[92] https://doi.org/10.1016/j.eswa.2013.03.030
[93] https://doi.org/10.1080/026999399379159
[94] https://doi.org/10.1080/02699938708408361
[95] https://doi.org/10.1038/s41562-018-0487-0

Lieberman, Matthew 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.

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[96]

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[97]

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[98]

MacKenzie, I. S. (2013). *Human-Computer Interaction: An Empirical Research perspective*. Waltham, MA: Morgan Kaufmann.

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–264.

Makel, M. C., Hodges, J., Cook, B. G., & Plucker, J. (2019). *Questionable and Open Research Practices in Education Research* (Preprint). EdArXiv.

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[99]

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*(2), 209–237. http://doi.org/10.1080/02699930802204677[100]

Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The Ability Model of Emotional Intelligence: Principles and Updates. *Emotion Review*, *8*(4), 290–300.

[96] `https://doi.org/10.1177/2372732216644450`
[97] `https://doi.org/10.1038/s41559-017-0160`
[98] `https://doi.org/10.3758/s13428-016-0809-y`
[99] `https://doi.org/10.1016/j.iheduc.2011.10.001`
[100] `https://doi.org/10.1080/02699930802204677`

http://doi.org/10.1177/1754073916639667[101]

Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge ; New York, NY: Cambridge University Press.

Mayo, D. G. (2020). P-Values on Trial: Selective Reporting of (Best Practice Guides Against) Selective Reporting. *Harvard Data Science Review*. http://doi.org/10.1162/99608f92.e2473f6a[102]

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[103]

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (2nd ed.). Boca Raton: Taylor and Francis, CRC Press.

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.2019.05.001[104]

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[105]

Molinari, G., 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, G., 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*

[101] `https://doi.org/10.1177/1754073916639667`
[102] `https://doi.org/10.1162/99608f92.e2473f6a`
[103] `https://doi.org/10.1016/j.ijer.2020.101589`
[104] `https://doi.org/10.1016/j.newideapsych.2019.05.001`
[105] `https://doi.org/10.1016/j.chb.2015.01.050`

Education, *24*(2), 129–144. http://doi.org/10.1007/BF03173006[106]

Molinari, G., 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[107]

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[108]

Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., ... Zwaan, R. A. (2016). The Peer Reviewers' Openness Initiative: Incentivizing open research practices through peer review. *Royal Society Open Science*, *3*(1), 150547. http://doi.org/10.1098/rsos.150547[109]

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[110]

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[111]

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[112]

Norman, D. A. (2013). *The Design of Everyday Things*. New York, NY: Basic Books.

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*,

[106] `https://doi.org/10.1007/BF03173006`
[107] `https://doi.org/10.1145/3004107.3004126`
[108] `https://doi.org/10.1037/a0033066`
[109] `https://doi.org/10.1098/rsos.150547`
[110] `https://doi.org/10.1016/B978-0-08-097086-8.25058-7`
[111] `https://doi.org/10.1016/j.ijer.2014.07.001`
[112] `https://doi.org/10.1016/j.jesp.2015.01.006`

348(6242), 1422–1425. http://doi.org/10.1126/science.aab2374[113]

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[114]

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). http://doi.org/10.1126/science.aac4716[115]

Osher, D., Kidron, Y., Brackett, M., 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[116]

Paquelin, D. (2011). La distance : Questions de proximités. *Distances Et Savoirs*, *9*(4), 565–590. http://doi.org/10.3166/ds.9.565-590[117]

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[118]

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[119]

Parkinson, B., Fischer, A., & 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[120]

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. http://doi.org/10.1037/0033-

[113] `https://doi.org/10.1126/science.aab2374`
[114] `https://doi.org/10.1037/1082-989X.8.4.434`
[115] `https://doi.org/10.1126/science.aac4716`
[116] `https://doi.org/10.3102/0091732X16673595`
[117] `https://doi.org/10.3166/ds.9.565-590`
[118] `https://doi.org/10.1016/j.chb.2007.05.006`
[119] `https://doi.org/10.1111/j.1751-9004.2011.00365.x`
[120] `https://doi.org/10.1177/1754073915590624`

2909.116.2.220[121]

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[122]

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[123]

Pekrun, R., & Linnenbrink-Garcia, L. (Eds.). (2014). *International Handbook of Emotions in Education*. 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., 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[124]

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[125]

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.

[121] `https://doi.org/10.1037/0033-2909.116.2.220`
[122] `https://doi.org/10.1016/j.learninstruc.2005.07.014`
[123] `https://doi.org/10.1007/s10648-006-9029-9`
[124] `https://doi.org/10.1080/02699931.2016.1204989`
[125] `https://doi.org/10.1016/S0191-8869(03)00084-9`

Pinker, S. (1990). A theory of graph comprehension. *Artificial Intelligence and the Future of Testing*, 73–126. http://doi.org/10.1145/2046684.2046699[126]

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[127]

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[128]

Posner, M. I. (2014). Orienting of attention: Then and now. *Quarterly Journal of Experimental Psychology*, 1–12. http://doi.org/10.1080/17470218.2014.937446[129]

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[130]

Reis, Rachel Carlos Duque, 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[131]

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[132]

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`

[126] `https://doi.org/10.1145/2046684.2046699`
[127] `https://doi.org/10.1037/bul0000026`
[128] `https://doi.org/10.4018/978-1-59140-562-7`
[129] `https://doi.org/10.1080/17470218.2014.937446`
[130] `https://doi.org/10.1016/j.tics.2015.09.003`
[131] `https://doi.org/10.1016/j.compedu.2018.01.015`
[132] `https://doi.org/10.1109/ICALT.2015.95`

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[133]

Riordan, M. A., & Kreuz, R. J. (2010). Emotion encoding and interpretation in computer-mediated communication: Reasons for use. *Computers in Human Behavior*, *26*(6), 1667–1673. http://doi.org/10.1016/j.chb.2010.06.015[134]

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[135]

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. http://doi.org/10.1037/h0077714[136]

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[137]

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[138]

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. (2018). An Appraisal-Driven Componential Approach to the Emotional Brain. *Emotion Review*, 175407391876565. http://doi.org/10.1177/1754073918765653[139]

Sander, D., & Scherer, K. R. (Eds.). (2009). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford, UK: Oxford University Press.

[133] `https://doi.org/10.1177/1754073908097189`
[134] `https://doi.org/10.1016/j.chb.2010.06.015`
[135] `https://doi.org/10.1145/130893.952914`
[136] `https://doi.org/10.1037/h0077714`
[137] `https://doi.org/10.1177/0018720815578267`
[138] `https://doi.org/10.1016/S0962-1849(05)80058-7`
[139] `https://doi.org/10.1177/1754073918765653`

Sangin, M., 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[140]

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.

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2020). *Why hypothesis testers should spend less time testing hypotheses* (Preprint). PsyArXiv.

Scherer, K. R. (1982). Emotion as a process: Function, origin and regulation. *Social Science Information*, *21*(4-5), 555–570.

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. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, *23*(7), 1307–1351. http://doi.org/10.1080/02699930902928969[141]

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.

[140] `https://doi.org/10.1016/j.chb.2010.05.032`
[141] `https://doi.org/10.1080/02699930902928969`

http://doi.org/10.1080/02699930500305016[142]

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[143]

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., 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.

Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality- and emotion-specific skills? *Personality and Individual Differences*, *53*(1), 16–21. http://doi.org/10.1016/j.paid.2012.01.026[144]

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.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90–100. http://doi.org/10.1037/a0015108[145]

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.

Scollon, C. N., Kim-Prieto, C., & Scollon, C. N. (2003). Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Stud-*

[142] `https://doi.org/10.1080/02699930500305016`
[143] `https://doi.org/10.1080/02699931.2018.1481369`
[144] `https://doi.org/10.1016/j.paid.2012.01.026`
[145] `https://doi.org/10.1037/a0015108`

ies, *4*(1), 5–34. http://doi.org/10.1023/A:1023605205115[146]

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.3.022806.091415[147]

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.

Siegert, I., Bock, R., Vlasenko, B., Philippou-Hubner, D., & Wendemuth, A. (2011). Appropriate emotional labelling of non-acted speech using basic emotions, geneva emotion wheel and self assessment manikins. In *Proceedings - IEEE International Conference on Multimedia and Expo* (pp. 0–5). http://doi.org/10.1109/ICME.2011.6011929[148]

Siemer, M., Mauss, I., & Gross, J. J. (2007). Same situation-different emotions: How appraisals shape our emotions. *Emotion*, *7*(3), 592–600.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution. *SSRN Electronic Journal*. http://doi.org/10.2139/ssrn.2160588[149]

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, 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[150]

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*.

[146] `https://doi.org/10.1023/A:1023605205115`
[147] `https://doi.org/10.1146/annurev.clinpsy.3.022806.091415`
[148] `https://doi.org/10.1109/ICME.2011.6011929`
[149] `https://doi.org/10.2139/ssrn.2160588`
[150] `https://doi.org/10.1037/0022-3514.93.3.431`

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, *28*(5), 1738–1747. http://doi.org/10.1016/j.chb.2012.04.014[151]

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[152]

Tobii AB. (2015). Tobii Studio.

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[153]

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[154]

Tullis, T. S., & Albert, B. (2013). *Measuring the User Experience. Collecting, Analyzing, and Presenting Usability Metrics* (Second edi). Waltham, MA: Morgan Kaufmann.

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[155]

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[156]

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. (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

[151] https://doi.org/10.1016/j.chb.2012.04.014
[152] https://doi.org/10.1007/s11412-006-9660-y
[153] https://doi.org/10.1177/1754073917742706
[154] https://doi.org/10.1207/S15389286AJDE1603_2
[155] https://doi.org/10.1080/02699931.2015.1087972
[156] https://doi.org/10.1111/j.1467-8721.2009.01633.x

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[157]

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[158]

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[159]

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.

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[160]

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[161]

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, *13*(5), 803–810. http://doi.org/10.1037/a0033839[162]

[157] https://doi.org/10.1080/02699931.2015.1081349
[158] https://doi.org/10.1016/j.copsyc.2017.07.017
[159] https://doi.org/10.1080/10463283.2011.627192
[160] https://doi.org/10.1111/j.1467-6494.1991.tb00252.x
[161] https://doi.org/10.1016/j.chb.2015.05.007
[162] https://doi.org/10.1037/a0033839