The Gistr Platform

Sébastien Lerique*

Abstract

Here goes the abstract.

Introduction

Studies of online interactions that are part of the "born-digital" sub-field of Digital Humanities usually focus on activities for which online records are readily available (mostly, blogs and large online social networks with varying degrees of access to the data). In stark contrast to "in vitro" laboratory experiments, this "in vivo" approach studies human interactions in an ecological environment by analysing their digital traces.[1]

Yet despite the large amounts of data and computing power that this quantitative observation approach takes advantage of, it remains very difficult to ask questions linking cognitive and social levels of human interaction, and the revolution that was expected to happen in social science has yet to materialise. More precisely: the question, both theoretical and experimental, of how to study and link together cognition and what social science defines as *social* proper, is still very much open.

Since better tools afford better questions and better theory, we focus here on the experimental side of this problem which is in fact quite simple: online interaction studies are based on digital traces, but most researchers have no control over what data is recorded in those traces, or over the way the interaction is framed and defined by the technical system that mediates it. As a result, only a very specific class of questions can be asked by online interaction studies: those that use only the available data and focus on the existing interaction framing, excluding any question for which some information is missing and not inferable or available elsewhere, as well as questions that need to tinker with the interaction framing. The latter is a very high barrier to asking questions that link the cognitive and social levels of human interaction.

The purpose of this document is to briefly introduce the approach of Web and Smartphone experiments – a promising method for human interaction experiments – and the trade-offs it makes, and Gistr, the Web experiment I am currently developing as part of my PhD thesis to overcome the shortcomings of our previous study on sentence transformation in the blogosphere.

---

*Centre d'Analyse et de Mathématique Sociales (UMR 8557, EHESS, Paris), and Centre Marc Bloch (Berlin). Email: sebastien.lerique@normalesup.org.

[1] I use the terms *in vivo* and *in vitro* in analogy to their meaning in biology. As explained in the main text, *in vivo* refers to the study of human interactions in their ecological environment, be it through direct observation or through digital traces of those interactions. *In vitro* refers to the study in conditions defined by the researcher, usually in a laboratory experiment.

1 Web and Smartphone experiments

Complementary to studies using existing digital traces, Web and Smartphone experiments strike a different balance in the trade-offs of experimental work and seem very promising in addressing the problems outlined above.

Indeed, smartphones and browsers (both on desktop and mobile) have evolved into powerful, ubiquitous application environments for which one can relatively easily develop any kind of experiment involving text, graphics, and human interactions. At the cost of higher engineering and recruitment efforts, this approach gives the experiment designer full control over what data is collected and the way interactions are framed. Given the omnipresent nature of smartphones, such experiments can also be deeply embedded in everyday life which, as I explain below, opens even more possibilities for questions on social interaction.

1.1 The need for embedding

Any quantitative study relies on abstracting out details of particular cases by reducing (most often averaging) values in each dimension to a few indicators. Being able to render a precise view of the studied phenomenon then depends on being able to determine which are the right dimensions to describe it, and having access to them.

Embedding experiments in the everyday life of subjects gives access to dimensions that can be otherwise unavailable: through the use of smartphones, an experiment designer can trigger interactions with subjects (for instance asking questions) at any moment of the day, or have measures running while subjects are offline (using the ever-increasing number of sensors present on the devices), both of which are impossible with digital traces. Above all, embedding an experiment means getting greater access to context, which opens the possibility of understanding phenomena the way they are meant in the lives of subjects, and not only in the way they are construed by the experiment designer.

Daydreaming is an example experiment developed as a smartphone application with Vincent Adam, Mikaël Bastian, Jérôme Sackur, and Gislain Delaire, that took advantage of this embedding. The experiment, focused on our awareness of daily mind-wandering, would probe subjects during a month at random moments of the day to ask them if they were mind-wandering (and, if so, what were the qualities of their thoughts).[2] While our team spent over a year developing the application, it allowed asking questions related to ecological situations which cannot exist in laboratory or passive collection studies. Figure 1 shows a sample question asked to the subjects, and Figure 2 shows an example of the results produced at the end of the experiment (this particular screen shows the results for one subject; seeing their own results was part of the reward for subjects participating in the study).

1.2 Pros and Cons

Experiments in the browser and on smartphones make specific trade-offs which differ from most other methods. The most important points are as follows.

Pros

- *Control*: similar to laboratory experiments, complex designs are possible, and all the interactions the subjects can be involved in are defined by the experiment designer. This includes for instance the way in which the experiment is framed (e.g. as a game or a self-improvement aid, aside from being a scientific study) and, more importantly, the ways in which subjects can interact with each other through the experiment.

---

[2]See http://daydreaming-the-app.net/ for more details.

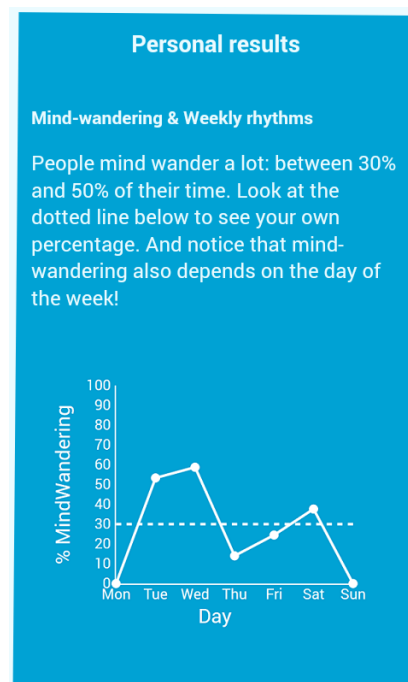Figure 1: Example questionnaire in the Daydreaming app

Figure 2: Results on weekly mind-wandering rhythms from the Daydreaming app

- *Embedding*: as explained above, smartphone-based experiments allow for real-life embedding: the experiment designer can choose when and how interactions with the experiment and between subjects take place, and measure any number of variables the device gives them access to (geolocation, time, phone agitation through its accelerometers, general noise level, etc.), virtually at any moment.
- *Scale*: if needed, the technical platform can relatively easily scale the number of subjects to several tens of thousands. This also applies to interactions between subjects, which can directly involve many people, be it at the same time or (for instance) in chains.
- *Flexible recruitment*: subject recruitment, while also a challenge (see *cons* below), is very flexible: services like Prolific Academic[3] let experiment designers recruit at reasonable costs in pools of tens of thousands of subjects with fine-grained demographic filters. Aiming for wider audiences can be done by offering rewards, framing the experiment as a self-improvement application, or turning it into a game.

Cons

- *Technical challenge*: developing such an experiment involves a substantial amount of engineering and makes use of several technologies most researchers, even technical, are not familiar with. While some all-in-one kits exist,[4] creating an experiment that meets one's research often requires learning average skills in most of the various technologies at play: a native or cross-platform smartphone environment, Web application development, backend server programming, and some server administration skills. Most importantly, the paradigms and problems encountered are often new to researchers: most programming is asynchronous due to network communication or user interface, and technicalities such as user management or email validation can grow into real engineering challenges.
- *Spam-control*: subjects are not constrained or encouraged by the face-to-face interaction of a laboratory experiment, neither are they (in most experiments) in the course of an interaction with friends where is it natural to them to not write spam, as can be the case in digital traces. Participants must have an incentive to perform the experiment's tasks well. If spam is naturally isolated in the experiment's design, one can for instance filter it once the data is collected and make payment depend on it prevalence. But if spam propagates in the experiment, real anti-spam pressures and motivations need to be factored into the whole design.
- *Recruitment cost*: while recruiting a few dozen or even a few hundred subjects is generally cheaper (not counting the cost of developing the experiment) than the equivalent for a laboratory experiment,[5] and can be very easy for fast prototyping and pilots, recruitment cost rises linearly with the number of subjects and the time they spend on the experiment, unless a different strategy is used. Turning an experiment into a playful application or an application useful to the user involves its own set of skills, can prove challenging, and must be factored into the development cost.

2   The Gistr Platform

2.1   Rationale

As part of my PhD thesis we are studying the transformation of short sentences – such as quotations from politicians or spokespeople – as they are propagated through various media. Our first study focused on the evolution of such short quotations as they are copied from blog to media outlet to blog. Indeed, authors often transform quotations when publishing them online despite the implicit and

---

[3]https://www.prolific.ac/

[4]See e.g. http://funf.org/ and http://www.epicollect.net/.

[5]Global competition on online platforms like Prolific Academic drive subject payments down.

common-sense injunction to quote people verbatim: a few words disappear, a contraction appears, the quote is cropped, etc. (Simmons, Adamic, and Adar 2011; Lauf, Valette, and Khouas 2013). Given this observation, the data collected by Leskovec, Backstrom, and Kleinberg (2009) is at first sight a very good candidate to study the evolution of online content as it is transformed by users. But the actual study proved itself much more challenging, for two fundamental reasons:

- *Missing information*: most blog and media outlet authors do not quote their source when they publish a quote online (it's often not relevant to the article), meaning there are no source-destination links in the data collected; this information must instead be inferred anew to study the evolution of content. There is also no access to author information (gender or age, experience in writing, but also psychological factors like memory span), ruling out any study of individual author effects in transforming the content.
- *Missing context*: the lack of access to the context of production and reception of quotes makes it impossible to interpret what a quotation means to its author or its reader (Wittgenstein 2010; Briggs 1992; Cuffari, Di Paolo, and De Jaegher 2014). Analysing any kind of semantic evolution is therefore out of reach for this kind of passively collected online data (Lerique 2016, to be published).

The Gistr platform[6] emerged from a concern to address these two problems. The general aim for this project is the study of interpretation and sense-making of short sentences in particular contexts, and the question of how interpretation and sense-making have global scale effects when accumulated and iterated.

2.2   State of the Art

[TODO: Primer on Epidemiology of Representations]

Up to now Epidemiology of Representations has focused on cultural bodily practices with long intergenerational lifecycles like

- smoking [Claidière & Sperber, 2007],
- bloodletting [Milton, Claidière, & Mercier, 2015],
- portraits [Morin, 2013], [TODO: read]
- religion [Boyer, 2001], [TODO: read]

interfacing with studies of diachronic evolution of language on the way [Claidière, Smith, Kirby, & Fagot, 2014]. [TODO: read]

Recently, practices with short intragenerational lifecycles like

- music [MacCallum, Mauch, Burt, & Leroi, 2012], [TODO: read]
- risk perception (abstracting away from individual variation) [Moussaïd, Brighton, & Gaissmaier, 2015 - in press], [TODO: read]

[TODO: add iterated learning tasks like Kirby et al.]

[TODO: revise ambitions and phrasing ('I')]

that have less to do with changes in bodily practice, and more with interpretation, have started to be studied. I aim to bring a new case in this area of short lifecycle opinion dynamics by studying the semantic evolution of short sentences and short stories in interpretation chains. What change takes place here is mainly due to interpretation and the reconstructional component of memory which, as mentioned above, involves many levels and is influenced by many factors. Therefore, after starting

---

[6]As explained further down, we aim for a series of experiments, based on the same paradigm, and gradually introducing complexity to the problem. Hence the term *platform*.

at the macro scale where individual variation and context details are abstracted out, I also aim to gradually move towards the meso scale, integrating more contextual and personal details and factors as I go along.

Finally, even if it's not my main target, this study should also be of interest to the big data/Twitter/Facebook studies community, where well-controlled data is hard to come by (if you don't work for Facebook).

2.3   General approach

Now interpretation is a kind of attention/perception, that is exploration of an environment. It's a process of exploring information (through educated attention) and putting it to use in the current task. This involves many different processing levels, and is studied by several different disciplines. Here are a few examples of what is involved:

- Interactions: relating to people involves several tasks; e.g. trying to look good to someone (based on what you believe that person thinks) is a task that will influence how you perceive what they say. For social anthropologists, meaning is a shared property that emerges from the interaction.
- Cognitive biases in general: for some tasks we often make recurrent and predictable mistakes or transformations (e.g. word frequency effect [Yonelinas, 2002], cognitive dissonance [Festinger, 1962], risk perception [Kahneman & Tversky, 1996] [TODO: read], and any number of cognitive-, memory-, or heuristics-induced biases), because of low-level processes involved in the task.
- Metaphors: reasoning and idea-association are closely related to basic metaphors that pervade everyday language (like 'more' is 'up', 'later' is 'forward', etc.).

2.4   Criticisms

Other disciplines work on this subject with different focuses and valuable criticisms. The three areas I roughly identify are:

2.4.1   Linguistics

I'm not a linguist, and this study is obviously affected by linguistics-level effects that I don't know about. I plan to address that issue, in a limited way, by augmenting my basic knowledge of linguistic effects by going to the LOT Summer School 2015, presenting my stuff at Martin Hilpert's discussion group "Issues in language variation and change", and asking around as much as I can.

2.4.2   Situated/embodied/extended cognition

Epidemiology of Representations rests on the fodorian idea of cognition where the brain is a storehouse of representations and everything happens in there. During the 2000's there has been a substantial move in philosophy of mind from this position to a position positing the central role of environment. In a nutshell, cognition is an exploration of the environment (vs. storing representations of the environment) and as such cannot be considered out of its environment. Some philosophers go so far as to consider that the environment is part of the cognitive device.

This line of thinking has greatly improved the way cognitive science analyzes contextual/situational information and interaction, has led to many experimental breakthroughs (that I know of, mostly in perception, e.g. perception substitution devices [O'Regan & Noë, 2001] [TODO: read]), and has a

number of implications for how perception and affect should be tackled (e.g. Bower & Gallagher, 2013 [TODO: read]).

2.4.3   Social anthropology

Adopting much of the "embodied" position w.r.t. representations, social anthropologist Tim Ingold criticizes Sperber's approach to cultural evolution on a number of points which can be summed up as follows:

- Culture is not something that is *transmitted*, but something that one *grows into* (Ingold, 1997). Hence the driving force of cultural change is not how representations are transmitted, but how attention is educated. As such, cultural and natural evolution are one and the same process, and separating them into one taking place before the other (Ingold, 2004) only hinders the analysis ("Cultural differences, in short, are not added on to a substrate of biological universals; rather they are themselves biological").
- Meaning (and therefore interpretation) is not a property of representations, it is what emerges from the individual's history with a particular representation; interpretation is how a piece of information is put to use in one's life (Briggs, 2006).
- An alternative to studying culture as an evolutionary process parallel to genetic evolution is to consider the whole ecosystem as a field of developmental systems (Ingold, 1998).

2.4.4   Treatment

Each of these disciplines would have many valuable criticisms addressing the shortcomings of how this study is currently conceptualized and planned. But as explained above, the plan is to start from the Epidemiology of Representations approach and, if time allows, gradually incorporate criticism as it serves the purpose of explaining the collected data and refining the conditions. (For instance, the embodied and social anthropology critique would prove useful in trying to develop conditions taking the context into account, which in turn would probably allow us to explain some noise.)

The bottom line is that all studies have their shortcomings and these criticisms will most likely not be addressed in the immediate future. Although the distant future sure would like to.

2.5   Breakout and development

2.5.1   Breakout

General goal: explore some interpretation effects, at the single and cumulative levels, in tasks involving sentence and story rewriting or reformulation. Interpretation involves many levels of complexity, so we start with the simplest possible condition (probably underspecified), and add new measures and conditions as we go to see if we can explain our noise.

To do this, we build an experiment where subjects repeatedly memorize and rewrite (i.e. interpret and reconstruct) textual content in various situations and tasks. This will be:

- Content we selected for the experiment
- Content newly suggested by the subjects
- Content already interpreted and transformed by previous subjects; this lets us generate interpretation chains, with each piece of initial content generating a tree of successive interpretations.

Measures [TODO: add references to state-of-the-art]

- How many (and which) elements are maintained (manually coded)
- LDA to see the evolution of subjects (either on each tree, or on the whole corpus)
- Author categorization based on their LDA bias and element creation/loss. Correlations to demographic measures (age, gender, CSP).

Perspectives

- Effect of stimulus degradation, e.g. change of direction in the bias [FIXME: ref]
- Effect of social proximity [Scherer & Cho, 2003; Binder, Scheufele, Brossard, & Gunther, 2011] [TODO: read]
- Effect of emotional state [Kramer, Guillory, & Hancock, 2014] [TODO: read]
- Effect of interaction setup/framing [Kahneman & Tversky, 1996, for framing] [FIXME: embodied ref for interaction?]

2.5.2 Development

The experiment presents itself as an online Game With a Purpose where people can freely register and participate. The gamification aspects are likely to come later, once the experimental aspects are stable and implemented. This lets us start experimenting on paid (Prolific Academic) or volunteer (Crowd Crafting) platforms, and later open up to platform-free participation.

Technically, it's a web application that runs in the browser and can also be packaged as a native app for Android and FirefoxOS. The backend to this app is spreadr: it receives, stores, and dispatches content to the app.

See the Milestones on GitHub for the detailed planned development.

References

Briggs, Jean L. 1992. 'Mazes of Meaning: How a Child and a Culture Create Each Other'. *New Directions for Child and Adolescent Development* 1992 (58): 25–49. doi:10.1002/cd.23219925804.

Cuffari, Elena Clare, Ezequiel Di Paolo, and Hanne De Jaegher. 2014. 'From Participatory Sense-Making to Language: There and Back Again - Springer'. doi:10.1007/s11097-014-9404-9.

Lauf, Aurelien, Mathieu Valette, and Leila Khouas. 2013. 'Analyzing Variation Patterns In Quotes Over Time'. *Research in Computing Science* 70: 223–32. http://www.micai.org/rcs/2013_70/Analyzing%20Variation%20Patterns%20In%20Quotes%20Over%20Time.html.

Leskovec, Jure, Lars Backstrom, and Jon Kleinberg. 2009. 'Meme-Tracking and the Dynamics of the News Cycle'. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 497–506. KDD '09. New York, NY, USA: ACM. doi:10.1145/1557019.1557077.

Simmons, Matthew P., Lada A. Adamic, and Eytan Adar. 2011. 'Memes Online: Extracted, Subtracted, Injected, and Recollected'. In *Fifth International AAAI Conference on Weblogs and Social Media*. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2836.

Wittgenstein, Ludwig. 2010. *Philosophical Investigations*. John Wiley & Sons.