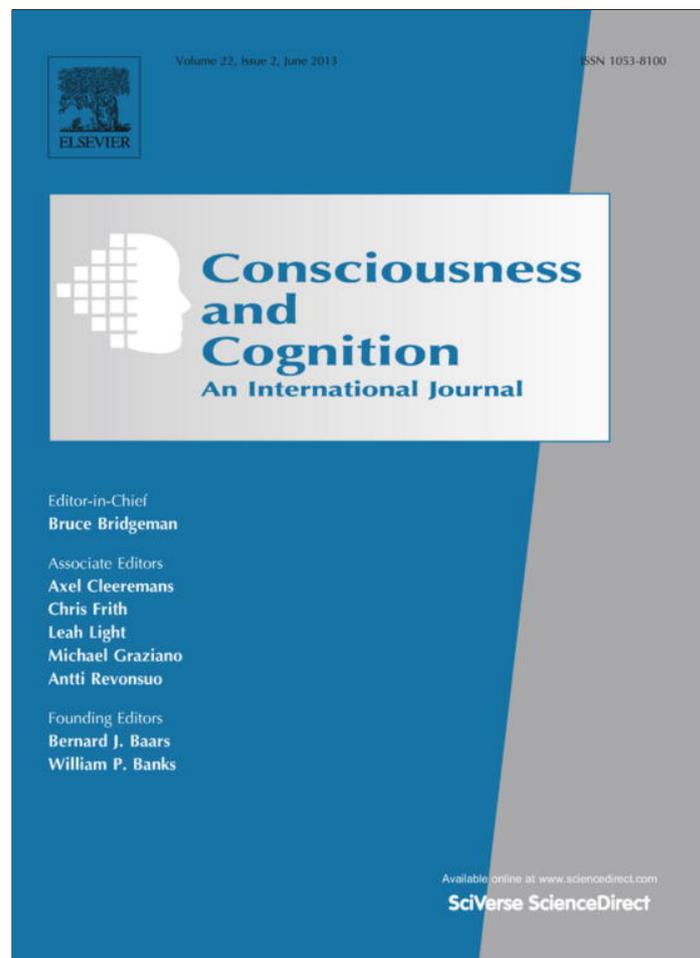


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/authorsrights>



Contents lists available at SciVerse ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

A gap in Nisbett and Wilson's findings? A first-person access to our cognitive processes



Claire Petitmengin^{a,d,*}, Anne Remillieux^a, Béatrice Cahour^b, Shirley Carter-Thomas^{a,c}

^a Institut Mines-Télécom (Télécom EM), 9, rue Charles Fourier, 91011 Evry, France

^b CNRS Institut Mines-Télécom (Télécom ParisTech), 46, rue Barrault, 75634 Paris Cedex 13, France

^c LATTICE (CNRS), 1 rue Maurice Arnoux, 92120 Montrouge, France

^d Archives Husserl (École Normale Supérieure), 45 rue d'Ulm, 75005 Paris, France

ARTICLE INFO

Keywords:

Awareness
Choice blindness
Decision-making process
Elicitation interview
Explicitation interview
First-person
Introspection
Lived experience
Pre-reflective experience
Passive memory

ABSTRACT

The well-known experiments of Nisbett and Wilson lead to the conclusion that we have no introspective access to our decision-making processes. Johansson et al. have recently developed an original protocol consisting in manipulating covertly the relationship between the subjects' intended choice and the outcome they were presented with: in 79.6% of cases, they do not detect the manipulation and provide an explanation of the choice they did *not* make, confirming the findings of Nisbett and Wilson. We have reproduced this protocol, while introducing for some choices an expert guidance to the description of this choice. The subjects who were assisted detected the manipulation in 80% of cases. Our experiment confirms Nisbett and Wilson's findings that we are usually unaware of our decision processes, but goes further by showing that we can access them through specific mental acts.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

One of the most quoted papers in the field of social sciences (almost 7000 citations in January 2013), written by Nisbett and Wilson (1977), draws from a series of experimental studies the conclusion that we have no introspective access to our cognitive processes, notably our decision-making processes: "There may be little or no direct access to higher order cognitive processes" (p. 231). The experiments of Nisbett and Wilson consist in manipulating the cause of the behavior of a subject, and then asking him to explain his behavior: most subjects provide an explanation without noticing the manipulation. The authors conclude that subjects "tell more than they can know" about themselves and describe processes that in fact they cannot access. This work has significantly contributed to the discredit of introspective methods, whose reliability and use in Cognitive Sciences is currently the object of an intense debate (Jack & Roepstorff, 2003, 2004; Overgaard, 2006; Petitmengin, 2009a, 2011; Varela & Shear, 1999; Weisberg, 2011).

A Swedish team of cognitive scientists recently tried to challenge the conclusions of Nisbett and Wilson (Johansson, Hall, Sikström, & Olsson, 2005) by carrying out the following experiment: the experimenter shows the participants two pictures of women's faces and asks them to choose which one they find the most attractive. Immediately after, he shows the chosen picture again and asks them to explain the reasons for their choice. But in some cases, the picture which is re-presented is the one that was not chosen (through the use of a double-card ploy, the subject does not realize the manipulation).

* Corresponding author. Fax: +33 1 60 76 42 86.

E-mail addresses: claire.petitmengin@telecom-em.eu (C. Petitmengin), anne.remillieux@free.fr (A. Remillieux), beatrice.cahour@telecom-paristech.fr (B. Cahour), shirley.thomas@telecom-em.eu (S. Carter-Thomas).

Surprisingly, the participants detect the substitution in only 27% of cases, and in 73% of cases, provide an explanation for the choice they did not make, a phenomenon that the authors term “choice blindness”.

Furthermore, a contrastive analysis of various linguistic markers (certainty, specificity, emotionality, complexity, etc.) shows no significant difference between verbal reports corresponding to manipulated and non-manipulated trials. This similarity leads the authors to conclude that even in the absence of manipulation, subjects verbalize what they think or believe to have done, not how they actually proceeded, which they are not aware of. The study was repeated a second time by the same team with similar results (Johansson, Hall, Sikström, Tärning, & Lind, 2006). It therefore seems to confirm the conclusion of Nisbett and Wilson that we have no introspective access to our decision processes. More specifically, this study casts doubt on the very possibility of remembering our past or just past experience, which is one of the cornerstones of introspection. As Stuart Mill had indeed noted, “a fact may be studied through the medium of memory, not at the very moment of our perceiving it, but the moment after: and this is really the mode in which our best knowledge of our intellectual acts is generally acquired.” (Mill, 1882/1961, p. 64). All the introspectionist schools of the early 20th century – that of Titchener (Titchener, 1910/1980), that of Binet in Paris (Binet, 1903), as well as the school of Würzburg in Germany – agreed on the fact that retrospection must be used each time the introspective act could disturb the target experience. However in the frequent cases where “introspection is retrospection” (Sully, 1881), how can we claim to introspect, for example the reasons for our decisions, if even the memory of the alternatives of the decision is so fragile that we can easily be fooled about them, and thus describe the reasons for a choice we did not make, without even noticing it?

Our practice of interview methods aimed at eliciting the disciplined description of the lived experience associated with a given cognitive process, led us to question the results of these studies. We reproduced the experiment of Johansson and colleagues, whilst introducing an “elicitation” stage for some choices. In the trials where subjects did not undergo an elicitation interview, our results are similar to those of Johansson and colleagues. However, in the trials where subjects were assisted in the elicitation of their experience, the substitution was detected in 80% of cases.

In a first part of this article we will describe our protocol and its differences with that of Johansson et al. After presenting our results in a second part, we will try in a third part to understand them: what did the subjects who were guided in the elicitation of their choice do that allowed them to detect the substitution, and that the others did not?

2. Protocols and methods

2.1. Johansson et al. protocols

In their first experiment (Johansson et al., 2005), Johansson and colleagues showed pairs of grayscale pictures of female faces to 120 participants (70 women) and asked them to choose which face in each pair they found most attractive. In some trials, immediately after their choice, the experimenter re-presented the chosen face and asked the participants to state the reasons behind their choice. Unbeknown to the participants, in certain trials, the experimenter used a double-card ploy to covertly exchange one face for the other, and instead presented the one that was not chosen.

Each subject completed a sequence of 15 choices, and was asked to justify his/her choice for 6 pairs, three of which were manipulated. The manipulated face pairs and the non-manipulated reports always appeared in the same position in the sequence. The deliberation time and the similarity of the face pairs varied. For time, three choice conditions were included: one with 2 s of deliberation time, one with 5 s, and one where participants could take as much time as they liked. For similarity, two sets of faces were created, a high-similarity (HS) and a low-similarity (LS) set.

Surprisingly, only 13% of the M trials were detected immediately. The subjects were then given different possibilities to detect the substitution retrospectively. All forms of detection combined, 26% of manipulated trials were detected. However, once a first detection was made, the participants became more suspicious, and more likely to detect future substitutions. To avoid such cascading detection effects, it was necessary to discard all trials after the first detection was made. This correction reduced the detection rate down to 8.2% (immediate detection) and 20.4% (all forms of detection). In 79.6% of cases, the subjects provided an explanation for the choice they did *not* make and never detected the manipulation.

Across all conditions, there were no significant differences in detection rate between the HS and the LS sets. There were no significant differences in detection rate between the 2 s and 5 s deliberation time conditions (20%, 16.6% with correction), but it was higher (36%, 28% with correction) when participants were given free deliberation time.

Johansson and colleagues reproduced this experiment with another group of 80 participants (49 women) (Johansson et al., 2006). The protocol was very similar to the previous one, except that a fixed deliberation time of 4 s was given to each participant, and the 15 face pairs were constructed in order to keep similarity constant at an intermediate level.

The overall detection rate for the manipulated trials was roughly equivalent to the results of the first experiment, with 27.5% of the trials detected. It therefore seemed to confirm the conclusion of Nisbett and Wilson that we have no introspective access to our decision processes.

2.2. Our protocol

2.2.1. The choice stage

We reproduced the experiment of Johansson et al. as closely as possible, but changed one condition, the way in which some of the verbal reports were gathered. As in the original experiment, each subject had to make 15 choices, and was asked

to justify his/her choice for 6 pairs, three of which were manipulated. We used exactly the same grayscale pictures as in the second experiment of Johansson et al. which had been designed in order to keep similarity constant at an intermediate level. We gave the choices which were followed by a verbal report the same position as in the initial experiment: positions 7, 10 and 14 for the manipulated trials, positions 5, 9 and 12 for the non-manipulated trials. We used a free deliberation time, a condition that Johansson et al. had identified as the most favorable to detection.

However among the three manipulated reports, two reports (which we will call “NEL” reports) were obtained in the same way as in the initial experiment, and the other through an “elicitation interview”¹ (“EL” report). At the beginning of the experiment, the subject is told that in some trials, he will be asked to describe his choice. In the case of NEL reports, the experimenter re-presents the chosen (or manipulated) face immediately after the choice, and asks: “I would like you to look at the picture carefully and to tell me why you chose it”. In the case of EL reports, immediately after the choice, the picture is given back to the subject, but face down. The experimenter then conducts an “elicitation interview” in order to get the subject to explain his choice process (see Fig. 1).

2.2.2. The elicitation stage

2.2.2.1. General presentation of the elicitation method. The “elicitation interview method” was initially developed by Vermersch (1994/2011, 1999, 2009, 2012) and then taken up in the context of cognitive science (Braboszcz, 2012; Petitmengin, 1999, 2001, 2006; Petitmengin et al., 2009; Valenzuela Moguillansky, 2012), as well as in the pedagogical (Maurel, 2009), managerial (Remillieux, 2010; Remillieux, Petitmengin, Ermine, & Blatter, 2010), technological (Cahour, Brassac, Vermersch, Pachoud, & Salembier, 2007; Cahour, Forzy, & Martin, 2010; Light, 2006), clinical (Petitmengin, Navarro, & Le Van Quyen, 2007) and therapeutical (Katz, 2011) fields. This interview method aims at helping a subject to leave the level of representations and beliefs in order to become aware of the way he really carries out a given cognitive process, and describe it with precision.

When asked to describe a given cognitive process, our natural tendency is to slip surreptitiously from the description of our actual experience toward the verbalization of justifications, beliefs, explanations, generalizations, and abstract knowledge about our experience. Contrary to what Dennett suggested by evoking our unchallengeable authority about our experience (Dennett, 2002; Schwitzgebel, 2007), and in spite of the puzzles raised by the concept of self-deception (Mele, 2001), our experience is not infallible, we can misinterpret the way it appears to us (Petitmengin & Bitbol, 2009). In the case which interests us – the decision-making process – numerous studies, of which the best known are those of Nisbett and Wilson, have indeed shown that when asked to explain the reasons for our choices and decisions, we tend rather to provide justifications or rationalizations, which substitute themselves for the actual decision criteria. This screening is even more efficient if it is “adaptative”, that is if it fulfils a function (Wilson, 2002), in other words provides a benefit. One function of justifications for our decisions is to convince our interlocutors of our merits, in order to try to strengthen our self-esteem. Another possible function is to reduce an internal conflict, which may be of a cognitive or an emotional nature. A cognitive conflict occurs when our ideas or beliefs are inconsistent with one another, triggering an uncomfortable feeling of tension called “cognitive dissonance” (Festinger, 1957; Stone & Cooper, 2001). A response to this discomfort may be to justify one’s action by inventing self-justifications. “The Fox and the Grapes”, one of the Aesop’s traditional fables, is sometimes used to illustrate the concept of cognitive dissonance: the fox that covets inaccessible grapes explains his departure and reduces cognitive dissonance by pretending that the grapes are too sour (Elster, 1983). Self-justification is notably used to justify unethical decisions, in order to make them more acceptable and maintain self-esteem (Mills, 1958; Myers & Bach, 1976). When the conflict is of emotional nature, justifications are rather dubbed “rationalizations”. Rationalization occurs “when the individual deals with emotional conflict or internal or external stressors by concealing the true motivations for his or her own thoughts, actions, or feelings through the elaboration of reassuring or self serving but incorrect explanations.” (DSM IV, 1994). This process has been studied in the psychoanalytic tradition, notably by Freud to account for the explanations given for neurotic symptoms (Freud, 1991), and by Fenichel who distinguished various types of rationalization (Fenichel, 1946).

To sum up, we often deceive ourselves about our decision criteria, and this deceit may have a specific function. Even when we try to “introspect” our decision criteria, we do not become aware of deceiving ourselves, because the confabulated criteria surreptitiously substitute for a precise retrieving of the actual decision experience.

However even when we are lucid about our decision *criteria*, something eludes us: our decision *processes*. For example, each time we dress in the morning, even if we know the reasons why we choose between several outfits (this one is more comfortable, more becoming, more suited to my activities of the day), we are largely unaware of the rapid internal operations we realize to make our choice (for example, make an inventory of my activities of the day, visualizing myself in this meeting with this outfit, imagining how my feet would feel after an hour’s walking firstly in these new shoes, and then in the old ones, etc.). We are seldom aware of how we proceed to choose, and only retain the outcome of this process, that is to say our choice criteria. This absorption into the object, the content, the “what” of our activity, to the detriment of the process, of the “how”, concerns not only our decision processes, but most of our cognitive processes. Whether we are touching, seeing, listening, imagining, remembering, understanding or deciding, whether we are performing a concrete or an abstract activity, a large part of our activity, although “lived through” subjectively, is not immediately accessible to reflective consciousness and verbal description. We experience it, but in an unrecognized or “pre-reflective” way (Depraz, Varela, &

¹ For the French “entretien d’explicitation”, which is sometimes also translated as “explicitation interview”.

Vermersch, 2003; Vermersch, 2000, 2009). The most surprising thing is that we are not aware of this deficit of awareness, which is the first obstacle in the way of becoming conscious of them: why should I make an effort to acquire an awareness I do not know I lack?

However this difficulty of access does not mean that our cognitive processes are out of reach. It means that accessing them requires a particular expertise, which must be acquired. Surprising as it may seem, we are blind to what is the closest to us, our lived experience, and we must learn to see it. This introspective expertise consists in carrying out specific acts (Depraz et al., 2003; Petitmengin, 2006, 2009b, 2011; Petitmengin & Bitbol, 2009). The “elicitation interview method” aims at triggering these acts, through specific prompts and questions, in order to help a subject to become aware of the unrecognized part of the process being described.

The first key to the elicitation interview consists of helping the subject to choose a particular occurrence of the cognitive process to be described, which is precisely situated in space and time, and bringing the subject back to this singular experience when he moves away from it towards the expression of comments, justifications, explanations and beliefs corresponding not to what he is experiencing but to what he thinks or imagines or believes about his lived experience (and thus interpreting it rather than describing it).

In many cases, there is a temporal gap between the initial experience and its description. The second key to the interview is thus to help the subject to retrieve or to “evoke” the experience, whether it is in the past or only just over (Vermersch, 2004/2011, Vermersch, 2009, Petitmengin, 2006). Evoking a past experience is a very specific process. It falls within a type of memory which Husserl called “passive memory” (Husserl, 1925/2001; Vermersch, 2004a, 2004b, 2006). We are always in the process of memorizing what we live, but often involuntarily, without being aware of memorizing. For example you did not voluntarily memorize the first thought you had when you woke up this morning. But you may be able to remember it. As the process of passive memory unfolds unbeknownst to us, we do not know what we know. This explains that at the beginning of an elicitation interview, people usually begin by saying “I do not know what I did, I do not remember anything.” However we can turn ourselves toward our past experience to make this information reappear. Interestingly, in passive memory the *recalling* of the memory also is often involuntary: it occurs spontaneously, usually through the intermediary of a sensorial trigger (Gusdorf, 1951). For example, to retrieve the first thought you had when you woke up this morning, you would probably have no way other than returning in thought to your bed at the moment when you awoke. Therefore in the course of an elicitation interview, even if the process to be explored is very fresh, because it has just been carried out, the interviewer helps the subject to retrieve precisely the visual, auditory, tactile and kinaesthetic, and possibly olfactory sensations associated with the very start of the process: “What were you seeing, hearing, feeling... at this moment?”. The subject “evokes” this moment when he recalls it to the point that the past situation becomes more present for him than the present situation is.

A set of objective – verbal, non-verbal and para-verbal – criteria for checking the effectiveness of the act of evocation have been identified (Vermersch, 2004/2011, Petitmengin, 2006). The verbal indicators are the use of the word “I”, specific context indicators of place and time, the concrete and detailed character (as opposed to conceptual and general) of the vocabulary used. The present tense is often a clue that the subject is re-enacting his experience, but it may also be a clue of generality (“usually, I do this”). An example of a non-verbal indicator is the direction of the eyes: the shifting and refocusing of the eyes, i.e. the fact that the subject drops eye contact with the interviewer and looks off into empty space, is a clue that the subject is accessing his inner world, ignoring the non-verbal reactions of the interviewer which could distract him from this task. At the same time, the flow of speech slows down, the words are often interspersed with silences, and co-verbal gestures often appear (Hendricks, 2009; Petitmengin, 2006): these clues show that the subject is retrieving his past experience and coming into contact with its pre-reflective dimension. It is only when, thanks to these clues, the interviewer verifies that the evocation state is sufficiently intense and stabilised that he can enable the interviewee, with the help of appropriate questioning, to turn his attention towards his inner processes and describe them. It is however rare for the interviewee to remain in the evocation state throughout the interview. Sometimes an ill-advised question or reformulation on the interviewer's part, or an external noise, can be sufficient for the interviewee to lose contact with the past experience. One technique enabling the interviewer to refresh the evocation state consists of formulating questions about the context of the past process, to which the subject cannot reply without “going back to it”. Another way is to ask questions about the content or the characteristics of the evocation during the interview, which the subject cannot answer without “refreshing” the evocation.

However evoking the experience is not sufficient to produce a comprehensive and detailed description of it. The third key to the interview consists of helping the subject to redirect his attention from the content, the “what” of his evoked cognitive activity, towards the involved process, the “how” (in our experiment, from the pictures and choice criteria towards the choice process). The *diachronic* dimension of the process corresponds to the stages of its unfolding in time: the succession of actions, perceptions and inner states which are lived. The *synchronic* dimension of the process corresponds to its configuration at a given moment in time, which cannot be described under the form of a succession: it includes in particular the type of attention and the sensorial modalities which are mobilized (Petitmengin, 2001, 2006). To collect these diachronic and synchronic descriptions, the interviewer asks questions which guide the interviewee's attention towards the various moments of the evoked process, without suggesting any content (Vermersch, 2004/2011, 2009). This type of “content-empty” questioning enables the researcher to obtain a precise description without imposing his own presuppositions. The structure of an interview is an iterative structure which consists of bringing the subject to evoke again his experience several times, while guiding his attention towards a diachronic or synchronic mesh which is finer each time, until the required level of detail is reached. It is very important to note that it is the reiteration of the act of evocation which ensures the reliability of the description.

2.2.2.2. *Use of the elicitation interview in our protocol.* In our protocol,² the interview is conducted as follows: the experimenter/interviewer starts by encouraging the subject to retrieve the moment where the pictures were presented to him: “*I suggest you go back a few seconds ago, to the moment when I showed you the pictures. To do this, I propose that you listen again to my voice when I ask the question “which of the two faces do you prefer?”, and that you see again the pictures as you saw them then.*” When the subject shows that he has arrived at this moment (a nod, a smile), the interviewer helps him to retrieve the different phases of his choice process, through questions that “point to” the different moments of the process, without inducing any content, such as: “*When you see the pictures, what happens? (...) What happens after this?*”. The same type of questions is repeated for each phase, in order to elicit a description of a finer and finer diachronic granularity. For example, if the subject says: “*First I look at the face on my right*”, the interviewer asks: “*When you look at the face on your right, what do you look at first? How do you go about looking at it? And then?*”. The interviewer draws the subject’s attention more specifically to the moment of the decision and the criteria of choice through questions such as: “*At the time when you finally pointed at this face, how did you know that you had made your choice? What had happened just before?*”.

Whenever the subject contributes new information, the interviewer helps him to deepen his description, through questions that draw his attention to the different – sensorial, attentional and emotional – dimensions of his experience at that particular moment, thus leading him to give a synchronic description, again without inducing the content of the responses. For example, if the subject says “*I started by looking at her eyes*”, the interviewer draws his attention to his mode of observation: “*At that time, when you are looking at her eyes, how do you do this? Are you only focused on the eyes, or do you perhaps see the whole face?*”.

Throughout the interview, it is the question “how” which guides the subject towards the description of more and more detailed elements of his evoked choice process. While in Johansson et al.’s protocol, the only question is: “*Why did you choose this picture?*”, in an elicitation interview, the question “why”, which deflects the subject’s attention to the description of explanations and abstract considerations, is never asked (Vermeresch, 2009). Rather, each time the subject escapes surreptitiously into explanations or generalities about his choice processes, the experimenter brings him back firmly to the evocation of the singular decision process he has just experienced.

Moreover regularly during the interview, and sometimes at the very end of the interview, the subject is invited to refresh the evocation of the moment of his/her choice through injunctions such as “*Take the time to retrieve this moment. . . these photos. . .*”, or through questions concerning the content or the characteristics of his evocation during the interview, for example for a visual evocation: “*What do you remember best in these photos?*”, “*While imagining these photos, where do you see them (at the top, at the bottom, to the right, to the left)? How far away are they? How big are they now?*”.

2.2.3. *The post-interview stage*

At the end of the interview (which lasted from 17 to 45 min), the interviewer invites the subject to turn over the picture and asks him if “*anything else concerning his choice process comes back to him when seeing this picture again.*”

In the case of explicit signs of detection during the second presentation of one of the manipulated pictures, the experiment is interrupted – in order to avoid cascading detection effects – and we go directly to the debriefing stage of the manipulation (see below). The trial is classified as “immediate detection” (ID).

In case of absence of any immediate signs of detection, as in Johansson and colleagues’ protocol, the subjects are given different possibilities to detect the substitutions retrospectively. At the very end of the experiment, the participants are asked the following questions: “*What did you think about the experiment?*” and “*Did you find anything odd about the experiment?*”. They are then explained the manipulation, and asked: “*Did you notice that some of the pictures I gave you back did not correspond with those you had chosen?*”. Then we ask them to indicate which faces they felt could have been manipulated, first without seeing them (which led them to indicate, for example: “*This was when you asked me a lot of questions,*” or “*This was the person who had earrings*”), then by picking up the manipulated picture(s) in the set.

Four levels of retrospective detection were defined according to the type of reaction of the subject. A retrospective detection of level 1 (RD1) corresponds to the case where just before or after the revelation of the trick, the subject spontaneously and immediately says that he noticed something unusual about the picture in question, and identifies and designates the picture clearly.

A retrospective detection of level 2 (RD2) is the case where after the revelation of the trick, the subject says he noticed something but does not evoke specifically the manipulated picture, and then goes on to indicate it in the set without any false positives,³ or with false positives, but indicated with a much lower degree of certainty.

In the cases of retrospective detection of levels 3 and 4 (RD3 and RD4), the subject says he noticed something and then finds the picture with false positives, or says he noticed nothing but finally finds it:

- without any false positive or with false positives he indicates with a much lower degree of certainty (RD3),
- with false positives and/or showing other significant clues of detection (during the experiment or during the post-test interview) (RD4).

² See Appendix A for examples of interviews, with and without elicitation. See Appendix B for the original interviews in French.

³ A “false positive” is a picture which is indicated although it was not one of those manipulated.

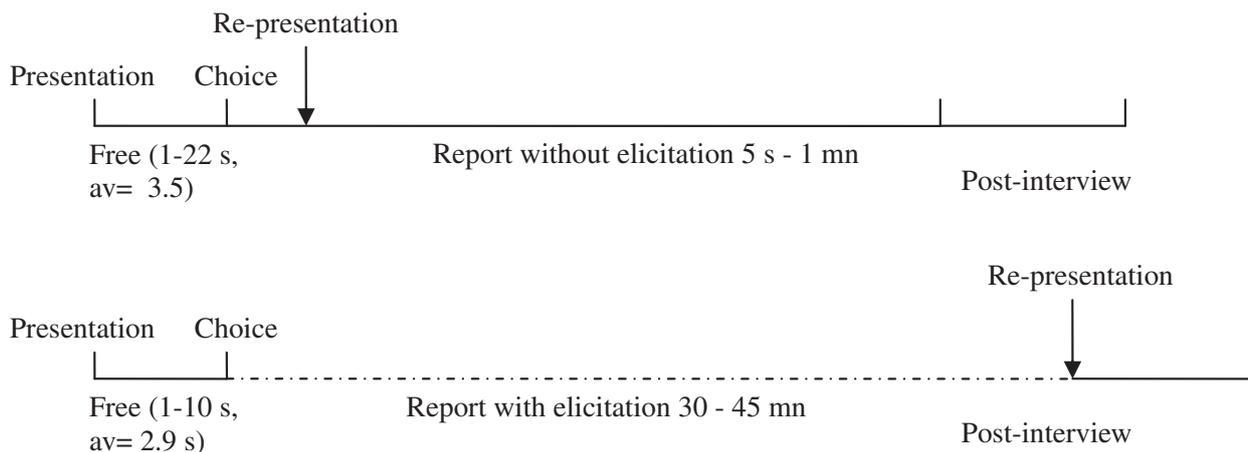


Fig. 1. Organisation of the trials with and without elicitation.

2.2.4. Additional details on our protocol

In order to homogenize the chances of detection of the substitution between the elicited and non-elicited choices, we implemented the following devices.

- From one interview to the other we varied the pairs of pictures presented, so that for each pair, the proportion of presentations was the same for the choices with and without the elicitation interview.
- We also varied the position of the choice which was followed by an elicitation interview (Nos. 7, 10 or 14 out of 15), in order to prevent, among other possible biases, the two following problems. On the one hand a “learning” bias: to do the interview systematically in first place might make the subject more attentive to his choice processes, which would increase the probability of detection for the following manipulated choices without elicitation. On the other hand a “trust” bias: if the elicitation interview always occurred last, the frequent absence of detection after the previous choices might lull the subject into a false sense of security which would reduce the probability of detection at the end of the elicitation interview.
- In the case of the NEL reports, when the experimenter re-presented the chosen (or manipulated) face immediately after the choice, he asked: “I would like you to look at the picture carefully and to tell me why you chose it”. We added this incentive to look carefully at the picture in order to prevent a possible attentional bias: having the elicitation interview before the second presentation might make the subject more attentive to the photo, so that there would be a greater probability of detection in this case.
- We considered the possibility of having the interviewer conduct the interview without knowing which cards were manipulated, but finally decided against this option. The main reason was that we wanted to reproduce the conditions of Johansson’s initial experiment as closely as possible, in order to evaluate the one condition which was modified, namely the type of interview. In Johansson’s experiment, the cards were manipulated and the interviews conducted by the same person. To realize a blind experiment, the manipulation of the cards and the interview would have to be done by two different persons, which would change the conditions of the trick and thus possibly the detection rate, in the NEL as well in the EL condition.

Moreover, we believe that a blind protocol, which assigns the functions of “interviewer” and “cards manipulator” to two different persons, would not be relevant for our experiment. If there is risk of induction, this risk concerns more the manipulator, who cannot for technical reasons ignore which test he is manipulating, than the interviewer. How in fact could the interviewer bias the results of the experiment during the interview? Obviously he cannot influence the choice of the picture, since this choice has been made before the interview. However could he incite the subject to detect the manipulation at the end of the interview? This could conceivably be done in two ways.

The first way would be to give the subject some indications about the content of the picture he actually chose. However we checked carefully that the interviewer, while encouraging the subject to remember his/her choice process, did not make any suggestions regarding the content of the choice (the faces), which would have biased the detection of the manipulation. It should be stressed too that taking care to avoid inducing the interviewee’s responses in terms of *content* – while helping him/her to realize the *act* of remembering – is a key principle of any elicitation interview. As all the interviews were recorded, the respect of this principle could easily be checked in the interview transcripts, such as the one which is provided in the annex.

The second way for the interviewer to induce the detection of the manipulation would be to communicate nonverbally during the interview his wish for the interviewees to detect the manipulation at the end of the interview. In fact, the interviewer’s aim in this case is bound to lead the interviewer to do everything to ensure that the interviewees remember their

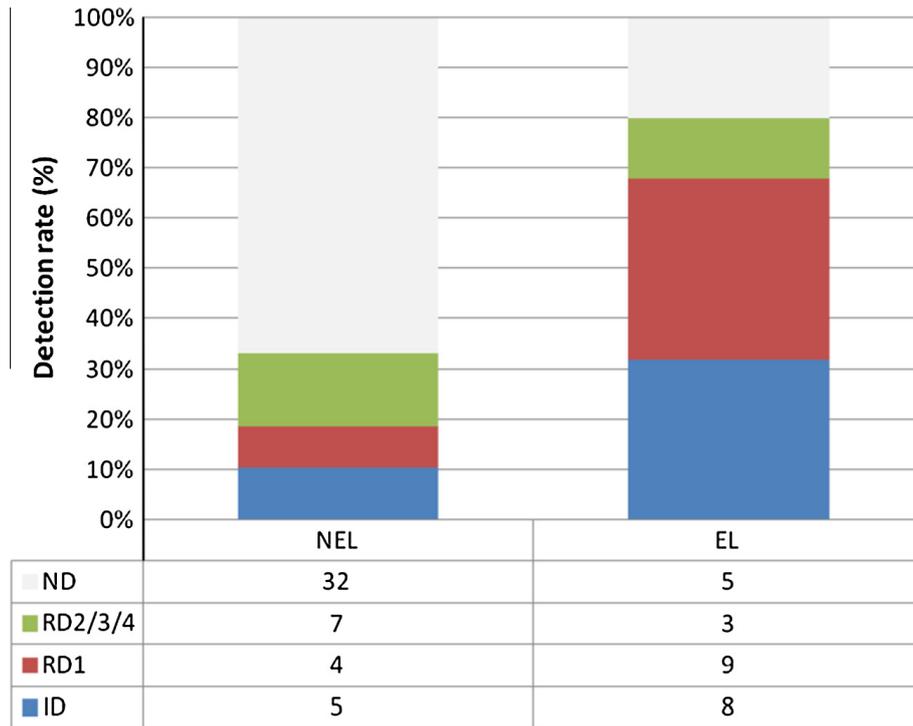


Fig. 2. Detection rate and frequencies with/without an elicitation interview.

choice process, which amounts to applying the elicitation interview method as defined in the experimental protocol as conscientiously as possible. In other words, in the specific case of our protocol, as the very goal of the interview is to trigger the acts which enable the detection of truth, it makes no sense to try to prevent this.

Thirty people (17 women) were interviewed according to this protocol. 141 reports were collected and analyzed, among which 70 “manipulated” reports, including 46 NEL reports and 24 EL reports. The interviewees were management or engineering students who were not trained in elicitation techniques.

3. Results

3.1. Quantitative analysis

3.1.1. Analysis of the detection rates

In this section, we first study detection rates depending on the EL/NEL condition (i.e. with or without an elicitation interview). The effect of the position of the trial on the detection rate is studied in a second stage. For comparison, global statistical indices (Chi-square and Cramer’s V^2) were used to highlight association between variables. We also drew up contingency tables to evaluate relative deviations (RDs) in each cell of the tables (see Bernard, 2003) in order to examine the weight of the different modalities of the variables. This second procedure is particularly well suited to comparing the different types of detection depending on experimental factors.

In Figs. 2 and 3, the upper part shows results as detection rates in order to allow a visual comparison as well as a comparison with the results of Johansson et al. (2005). The lower part shows the corresponding raw frequencies.

3.1.1.1. *Global dependency of detection type to elicitation interview.* The global statistical comparisons indicate a powerful impact of the EL/NEL condition on the detection outcome ($\text{Chi}^2(3) = 18.50; p < .001; V^2 = 0.35$). As presented in Fig. 2, when taking into account all types of detection, 80% (20/25) of the EL trials resulted in detection,⁴ against only 33% (16/48) for the NEL trials. In this respect, for NEL trials, the detection rate was close to that obtained by Johansson et al. under comparable conditions: 36% (28% corrected) in the first experiment with the “free deliberation time” condition (Johansson et al., 2005) and 27.5% in the second experiment (Johansson et al., 2006).

3.1.1.2. *Details on detection types based on relative deviations.* Contingency tables enable an estimate to be made of the local association between modalities of variables with relative deviations (RDev). RDevs are calculated by comparing each cell of the table – the observed frequencies – with calculated expected frequencies (i.e. those would have been obtained if there was

⁴ The statistical significance of this outcome has been tested by the p -value method, with a null-hypothesis 1/3 for the probability of detection. The calculated p -value for obtaining 20 detections or more in a sample of 24 is 7×10^{-7} . The result can then be considered as highly significant.

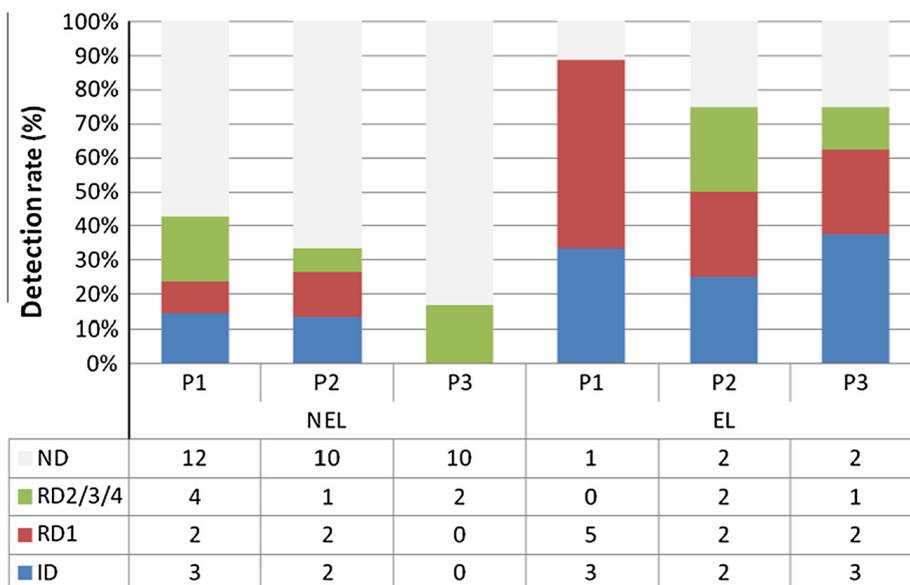


Fig. 3. Detection rates according to type and position of the trials.

Table 1

Contingency table of relative deviations for detection type and EL/NEL condition.

	ID	RD1	RD2/3/4	ND
NEL	-0.415	-0.532	0.065	0.315
EL	0.797	1.022	-0.124	-0.605

no association between the variables). The association is positive (i.e. there is an attraction) when the RDev is positive; and negative (repulsion) otherwise. By convention, we report only attractions with values >0.20. RDevs highlight which modality of a variable (v1), tends to be more associated to which modality of another variable (v2).

As shown in Table 1, the ID rate is particularly significant for EL trials (32%, RDev = 0.797) and it is specifically weak for NEL trials (10%, RDev = -0.415). The same deviation is observed for retrospective detection of level 1 (RD1) with a strong frequency of RD1 for EL trials (36%, RDev = 1.022) and a weak frequency for NEL trials (9%, RDev = -0.532). Conversely, the ND rate is small for EL trials (20%, RDev = -0.605) and large for NEL trials (77%, RDev = 0.315). But in the RD2/RD3/RD4 condition, the detection rates are not significantly different (12%, RDev = 0.065 for EL trials and 14%, RDev = -0.124 for NEL trials). These detailed comparisons indicate that the differences are more substantial in the case of earlier signs of detection rather than late signs.

3.1.1.3. *Effect of the position of trials.* The position of the trial also had an effect on the detection outcome with regard to the EL/NEL factor ($\chi^2(15) = 27.78; p < .05; \eta^2 = 0.30$). Fig. 3 and Table 2 show the differential effect of EL and NEL trials on detection rates depending on their position. For EL trials, position 1 (P1, trial no. 7) specifically enhanced ID (RDev = 0.872) and RD1 (RDev = 2.120), with no RD2/3/4 and only one ND. For NEL trials, it is position 3 (P3, trial no. 14) that is different from the two others, with no ID and no RD1 (for both: RDev = -1). In other words, the elicitation interview enhances earlier signs of detection at the beginning of the experiment whereas the absence of an elicitation interview makes detection very rare at the end of the experiment. We will come back to these results in Section 4.

3.1.2. *Analysis of the deliberation times*

We also conducted an analysis of the deliberation times⁵ for the 72 manipulated choices. The results are the following:

- The average deliberation time is 3.3 s.
- If we take all the trials into account, the detection rate is higher for longer deliberation times: the average deliberation time is 6 s for the (13) choices which are followed by an immediate detection, and 4 s if we take all types of detection into account (36 choices), whereas it is only 3 s for the (37) choices which are not followed by detection. These results are in line with those of Johansson et al. who found a higher detection rate when participants were given free deliberation time (Johansson et al., 2005).

⁵ To compare means, we should have the same number of people in each group, and at least twenty people in each group, which is not the case. Nevertheless the comparisons revealed some interesting results.

Table 2

Contingency table of relative deviations for detection type, EL/NEL and interview position.

		ID	RD1	RD2/3/4	ND
NEL	P1	−0.198	−0.465	0.390	0.127
	P2	−0.251	−0.251	−0.513	0.315
	P3	−1.000	−1.000	0.217	0.644
EL	P1	0.872	2.120	−1.000	−0.781
	P2	0.404	0.404	0.825	−0.507
	P3	1.106	0.404	−0.088	−0.507

P = position of the manipulated trial in the sequence of choices.

P1 = position 1 (trial no. 7), P2 = position 2 (trial no. 10), P3 = position 3 (trial no. 14).

- Although the conditions of deliberation are strictly identical, in the EL trials the average decision time (2.9 s) is lower than in the NEL trials (3.5 s). Furthermore the average decision time of the choices which are followed by an immediate detection is 3.1 s in the EL trials, and 10.8 s in the NEL trials. These differences in average deliberation times between EL and NEL trials are explained by the presence of two NEL trials with particularly long decision times (16 s and 22 s), which can only be explained by chance. If we do not include these two trials, the average decision time in the NEL trials is 2.8 s, that is very similar to the average decision time in the EL trials (2.9 s). And the average decision time in the NEL trials that led to immediate detections is 5.3 s, which means that the only NEL trials which led to immediate detections (5 out of 48) started with long deliberation times.
- Deliberation times for NEL trials preceding and following the EL trial are identical (3.5 s).

We will come back to these results in Section 4.

3.2. Comparison of the verbal reports with and without elicitation

We conducted a detailed comparative analysis of both types of reports, EL and NEL.

The first finding is that the EL reports are much longer than the NEL reports: 2900 words versus 208 words on average. The analysis then shows that the NEL reports are focused on the criteria of choice: facial features and/or feeling aroused (“the one on the right is smiling,” “she has more delicate features,” “I find her friendly”). On the other hand, the EL reports contain, in addition to the *criteria* of choice, descriptions of the observation and choice *processes*. Subjects describe *how they went about* watching the faces, comparing them and deciding. For example: “I look at her globally, then I zoom on the mouth”, “I imagine how she would react if I said hello”, “I compare by keeping the first face in mind and superimposing it on the second like a tracing”.

A preliminary linguistic analysis of EL and NEL reports reveals some significant differences which tend to confirm these observations. Pronoun distribution, for example, is very different. The proportion of verbs which are introduced by the pronoun “I” is proportionally twice as high in EL reports (60 versus 28 per 1000 words) reflecting the fact that the descriptions are primarily focused on the subjects’ own choice processes. Conversely the number of occurrences of the pronouns “she” and “they” is almost twice as high in the NEL reports (21 versus 12.5 per 1000 words), in which the descriptions are essentially focused on the characteristics of the faces. The reflexive character of EL reports is further confirmed by the much greater number of reflexive verbs, indicating a switch in the subjects’ attention towards their internal processes (6.1 versus 1.7 per 1000 words): “je me dis” (I say to myself), “je me demande” (I ask myself), “je m’assure” (I make sure),⁶ “je me focalise” (I focus my attention), “je m’imprègne” (I let myself become permeated by”). The action-focused character of the descriptions is also confirmed by the greater variety of verbs in the EL as opposed to the NEL reports (104 versus 33 different verbs), with nearly all the 71 extra verbs in the EL reports being action verbs (for example: I start, I stay, I go back, I stop, I look for, I compare, I keep, I stay). Interestingly, the French verb “faire” (to do/make/perform an action) does not appear even once in the NEL while it appears 73 times in the EL, for example: “je fais des comparaisons” (I make comparisons), “je me fais une image dans ma tête” (I make a picture in my head), “je fais des aller-retours (I go back and forth).

We also performed an analysis of the content of the EL reports. Such an analysis aims to identify the *structure* of the experiences described, that is a network of relationships between descriptive categories, independently of the experiential content, and to detect any generic structures which are gradually extracted from the initial descriptions thanks to a succession of abstraction operations (Petitmengin, 1999; Petitmengin & Bitbol, 2009). This enabled us to identify three types of choice strategy: immediate or non-immediate, positive or negative, “features-focused” or “feeling-focused” strategy.

- (1) The “immediate” choice (9 trials/20⁷ i.e. 45%) is made easily, whereas in the “non-immediate” choice (11 trials/20 i.e. 55%), the subject has trouble deciding. This subjective evaluation of the difficulty of choice and decision time is consistent with the objective decision time (an average of 1.2 s for the immediate choices and 4.5 s for the non-immediate choices).

⁶ In some cases the English equivalent does not require the reflexive pronoun to indicate a reflexive meaning.

⁷ The trials which were difficult to classify according to a particular strategy were not taken into account. This explains why the total number of trials is sometimes inferior to 24.

Table 3
Positive and negative strategies criteria.

	“Features” criteria	“Feeling” criteria
“Positive” strategy	<p>“The face is smoother, finer features, more pleasant and regular”</p> <p>“The hair and hair color ... the gaze ... the nose and mouth, and the face shape as well”</p> <p>“She really had a little smile ... that’s what has reinforced my choice”</p> <p>“There is this lock of hair on the side ... which I liked. It was more natural, less overdone ... the lock of blond hair”</p>	<p>“Maybe I felt closer to her. I am slim too and so ... since she is also slim ... I say to myself we have something in common. I don’t say it too much in fact, it’s more an intuition. It’s more a feeling, more of a general state of my body ... I feel ... slim. I saw it was with her that ... I have more in common so I chose her”</p> <p>“I wondered who was the one I felt the closest to, who I would have more confidence in. Perhaps also the person who I would most likely go and talk to, and exchange with afterwards, who would reassure me the most”</p>
“Negative” strategy	<p>“The choice was made on the eyes. It is after having noticed ... well, the irregularity of the eyes of the person ... that I did not choose her. I thought that I would choose this one”</p>	<p>This [selected] faced shocked me, appealed to me less than the other. I chose it because the other made me think, well ... made me think of someone else, in fact, I chose this one by eliminating the other”</p>

- (2) The subject uses a positive strategy (9 trials/18 i.e. 50%) when he reaches a decision based on the face he prefers, and a negative strategy (9 trials/18 i.e. 50%) when he decides on the basis of the eliminated face. The (positive or negative) choice criteria are either the features of the face or the feeling aroused by it (see Table 3).
- (3) The **“features-focused” strategy** (5 trials/22 i.e. 22.7%) consists in exploring and comparing the features of the faces, the way we play the “spot the seven differences game” (a metaphor that appeared in four interviews).

“One remembers the look, the shape of the nose, of the face, that we have seen, and one tries to see the similarities and differences between the face we have just seen and the face we see now, to compare, after say ‘well, as a result, there, which of the two, which one I preferred.’”

Some participants describe a chronology of exploration of the features⁸:

“The eyes, the position of the nose, the mouth, and then the limit of the chin.”

“If I had to represent the course of my gaze, it was the hair, then it went down uh ... the opposite side from the kind of strand of hair she had, then went up, yes, that’s it. [The way I went round the face] is precisely in a clockwise direction ...”

“There is the quick look, then there is the hair, then I look at the hair of the picture next, and then I go down, yeah, and so on.”

The mode of attention is visual and focused: the features are “scrutinized” one by one.

“Once one has seen that the nose was not straight bah one will zoom. I look at her nose ... well I finally confirm that she has something wrong.”

“I see the mouth and also that there are small wrinkles, expression wrinkles. The rest I don’t see it very well. I’m only interested in the smile.”

“It’s as if I was zooming on the mouth.”

The comparison is carried out either by rapid back and forth movements between the two pictures, or by “superimposing” the two faces (see Table 4).

In the latter case (“superimposing” the two faces), the comparison requires a “visual retention”, that is to say the pre-reflective and involuntary creation of a more or less complete mental image of at least one of the two faces, which is superimposed on the picture being observed.

“The visual image remains in the short term, remains imprinted ... temporarily.”

“I had the image of the previous one in my head.”

“I keep it in memory. It remains in my head. I have her in my mind but it is a little fuzzy compared to the picture I am watching. It is fuzzy compared to the picture I have in front of my eyes.”

On the other hand, the **“feeling-focused” strategy** (6 trials/22 i.e. 27.3%) consists in evaluating successively not the features of the faces but the feeling or emotion aroused by each, and then comparing them.

“I had a general impression. The first thing, was the emotion, well the feeling.”

“I look and it will inspire something to me. I need to have a sensation, to feel something before [comparing]. [I look at the picture] more for the sensation. It’s more the look that emanates from her when one looks at her, is not even really looking at her, it’s something which happens when you look at her, how you feel the person.”

⁸ A line of objective validation of this type of description would be to compare them with recordings of eye movements: when the subject says he starts by looking at the picture on the left, does he really proceed this way?

Table 4
“Features-focused” strategy.

Alternative observation of the faces	<p>“I go back and forth between the two pictures”</p> <p>“There it’s really looking at the details, then coming to the other one and ... it’s going back and forth, in fact, between the two photos”</p> <p>“I juggled between the two pictures ... I looked at the hair, comparing with the other, it’s a checklist: that, it’s done. That’s done... and I go to the next point and so on”</p> <p>“In order to scan all the criteria that I may have, the differences, as in the differences game a bit, when you’re a child and you play the differences game, so I try to play the differences game. I think I look at them in turn, quickly. I don’t think I look at them at the same time ... to find differences I don’t look at the same time, I look on the left, then on the right, for example, is this one tired, has she dark rings under her eyes? And the other girl, has she dark rings under her eyes? Is she wearing makeup? And the other girl, has she got makeup on?”</p> <p>“First I will look at one picture and then another, well, I’ll return from one photo to the other ... very quickly we’ll say”</p>
Mental superimposition of the faces	<p>“They were linked in my memory like that, to make the decision, like a tracing. Inside my head”</p> <p>“And I compare while keeping the first one in memory. It is superimposed as a layer. I have the picture in my head and I visualize the picture that I’m looking at and the one I remember, it appears like with the tracings kids in kindergarten do and so I superimpose a layer and I look at the points, the seven differences game”</p> <p>“It was... between brackets, two images that I superimposed. It’s like taking a tracing paper and saying ... well, I have drawn the main features and ... if I superimpose the two tracings, well, it works”</p> <p>“It’s easier to compare when they are one upon the other and this really allows me to have both together and as a result it’s easier to compare. [I really put them] one upon the other, that is to say that I will, for example, put the eyes on the eyes, and now ... really on the hand [location where the picture has been observed]”</p> <p>“It allowed me to see that it was thinner, because in the juxtaposition the other face was wider, its outline went over that of the first face”</p>

Table 5
“Feeling-focused” strategy.

Perception of the face as a person (as opposed to its perception as a photo)	<p>“I had the impression she was wearing a hooded sweatshirt, like me, her hair was pulled back as if she was playing sports, not necessarily made up or anything. She had a hood. I thought, but I thought, well, very, very subtly that she was a sporty person, one can feel that she’s a sporty person”</p> <p>“I try to imagine approximately her corpulence. I imagined, with the shape of her face, it was a somewhat corpulent person. But ... with a round face, quite pleasant. I try to imagine approximately her style of dress. Then, I imagined her clothes ... not ample, but quite feminine, but ... not tight, pretty ... I thought they went well with her face and hair”</p> <p>“The moment I look at her eyes, I rather see a person, since I think she is sad... it reminds me of feelings and therefore, she is a person. I’ll try to understand why she feels that way. I try to put myself in her situation. Why, me, would I feel like that. I try to go back to memories of ... events that may have affected me or made me ... like her”</p>
Contextualisation of the person	<p>“I see them at work, yes ... well I see them in the hall of offices, a bit like here in fact”</p> <p>“I imagine her queuing at the bakery, looking less made-up a little as if she had just jumped out of bed on Saturday morning”</p> <p>“I just imagined them in the subway”</p> <p>“I would imagine them well in the street, with people wandering in a street, rather a busy street with ... activity. Not a wide boulevard, a large avenue, but a medium-sized street in Paris”</p>
Imagination of an interaction	<p>“I imagine would I go towards her easily. I imagine myself talking to her, easily. I imagine actually myself being in front of her and her answering me. It happens, there is an exchange. She seems to look at me and she’s not cold. I feel there may be an interaction”</p> <p>“(I see) something angular and a little hard, something closed, which doesn’t let me come in. Something that is not favorable to a relationship or exchange or circulation”</p> <p>“(I feel) I step back, I shrink back. It’s not disgust, (but) a step backwards, (in) the legs, a step back, the right foot which wants to step back, not the left but the right”</p> <p>“I imagine how she would react, in fact, if I did that, if for example I said hello, what would she, how would she react, how would she smile”</p>

To do this, the photograph is “personified” (see Table 5):

- the face is put into context, that is to say imagined as part of a body and sometimes in a specific location;
- an interaction, or rather the hint of a possible interaction, for example an exchange of glances, is imagined;
- the reaction (the desire to talk to the person, or on the contrary a feeling of unease or fear) is evaluated.

This strategy is associated with a mode of attention which is global, defocused: the face or the person is observed as a whole; and receptive: it is a matter of letting oneself be permeated by what “radiates” from the person.

“I looked globally. One sees the whole face ... I can still describe it, it was quite round, she looked blonde, light eyes, I see the entire photo. I saw actually the picture as a whole.”

“I look at the overall appearance . . . in a holistic manner.”

“I’m not going to focus on anything in particular but really on the whole and what emanates from it. I see her, the global expression.”

“I start with the overall look, well what emanates from the photo, well from the person.”

“I see the face in the background, I am focused on the smile but I see the face in the background. It’s not as if the smile took all my vision, I see in the periphery the rest of the face and of the picture.”

“Retention” of one of the faces may be used during the observation of the other, but in this case it is not a visual retention but a retention of the feeling which is aroused (a “felt retention”).

“It comes back with the overall impression I might have of the person before. I already have an a priori about the person I looked at. Basically it’s as if I could still hear a little voice in my head telling me, ‘yes, but with the other you got a different impression’, it comes to confirm what I might have felt before. It’s not a voice, it’s more a kind of feeling, an intuition. Not even in my head actually. That would be there . . . in my heart.”

In 11 trials out of 22 i.e. 50 %, the subject uses a **mixed strategy**, combining a features-focused and a feeling-focused strategy. All these strategies sometimes include an internal criterion informing the subject of the fact that his decision has been made (sense of relief, of determination. . .).

“I feel better. I feel that this may be a relief. (. . .) It is something comfortable in fact, the feeling of: ‘it suits me’. It calms me down in fact. (. . .) Even the way of breathing . . . I have this feeling . . . that I breathe . . . well that I breathe out more (easily) in fact.”

“Determination, (. . .) I am convinced, it is: “That’s it, I know it, that’s all. It is this one”. (. . .) And I do not regret it later.”

The analysis of the interviews also revealed a variation in the detection rate depending on the type of strategy used, an additional result to which we will return to in Section 4.

4. Discussion

In summary, without elicitation, the subjects only detect the substitution in 33% of the cases, while the picture remains hidden to them for 2 s. At the end of an elicitation interview, the detection rate is 80% while the picture remains hidden during the whole of the interview, that is 17–45 min. How can this discrepancy be explained? What do the subjects do in this case, which explains why they detect the substitution much more frequently, while the time between the initial observation of the photo and the presentation of the trick photo is much longer, which should greatly reduce the chances of detection?

This experiment was especially designed so as not to trigger a specific effort of memorization at the moment of the choice, through particular instructions aimed at making the events more salient at that time (Nisbett & Wilson, 1977, p. 251). As Johansson et al. explain, “this experiment was meant to simulate a choice situation in which no prior evidence indicates that a high level of monitoring is needed” (Johansson et al., 2005, Supporting Online Material p. 4 and 10). It corresponds to situations of everyday life where we do not voluntarily memorize the alternatives nor the reasons for our choices, the situations that are the most common.

The conditions of memorization in the EL and NEL trials are identical in this respect. The difference in detection rates cannot therefore be explained by a greater stimulus salience at the moment of the choice in the EL condition, which would explain better memorization, and consequently slower memory decay than in the NEL condition. The subject never knows if the choice he is making will be followed by an elicitation interview or not. Moreover as we have shown in Section 3.1.2, the average deliberation time is lower in EL trials (2.9 s) than in NEL trials (3.5 s), and in trials which result in immediate detections (IDs) it is more than three times shorter in the EL trials (3.1 s) than in the NEL trials (10.8 s). Thus the higher detection rate in the EL trials cannot be explained by a special effort of memorization at the moment of the choice in anticipation of the interview, which would have resulted in a longer decision time.

In other words the higher detection rate cannot be explained by slower memory decay. On the contrary, the fact that the delay between the two presentations of the pictures is much longer in the EL condition would predict a much lower detection rate – the longer the delay the lower the detection rate. The fact that this rate is much higher, which cannot be accounted for by a difference in the *memorization* of the faces at the moment of the choice, would only seem to be explained by a specific and efficient act of *remembering* at the time of the report.

If we go on to consider the detection rates according to the position of the trials (Fig. 3), we observe that they are higher in the first manipulated trial than in the second and third trials, in the EL condition as well as in the NEL one. This decrease in detection rates between the first and the second trial shows that the experiment did not trigger a “learning” effect: the interviews did not generate a greater effort of memorization in the subsequent trials, which would have resulted in higher and higher detection rates in the course of the experiment. The absence of learning effect is confirmed by the fact that the average deliberation time remains stable throughout the experiment. However this decrease in detection rates may be explained by an interference effect (Peterson & Peterson, 1959): the increasing number of faces which have been seen in the course of the experiment may make it more and more difficult to remember the two faces which have just been seen. The earlier detection in the EL trials as well as the higher detection rates in both types of trials in the first position, could be explained by the absence of this effect at the beginning of the experiment. But while the detection rate decreases dramatically in the

third NEL trial, it remains stable in the second and third EL trials. This continuous decrease in the NEL trials could be explained by a “trust” effect, in addition to the interference effect: the absence of detection in the two first trials may have generated a sense of security, which resulted in a sharp fall of the detection rate. However the trust effect, as well as the interference effect, do not occur in the EL condition, indicating that they are counterbalanced by a specific activity during the interview.

These observations lead us to conclude that the difference in detection rates between the EL and NEL trials cannot be explained by a difference in the memorization of the faces at the moment of choice, but by differences between the EL and NEL conditions at the time of the report.

The main differences between the results in the EL and NEL conditions are the following:

- (1) higher detection rate in the EL condition;
- (2) far more detailed descriptions of the choice processes in the EL condition.

The main differences between the conditions of the reports are the following:

- (a) in the EL condition, subjects are periodically asked to recall the pictures and the moment of the choice;
- (b) in the EL condition, more questions are asked.

From these conditions and results, we draw the following conclusions:

(A): (1) is due to (a). It is the fact that participants are periodically asked to recall the pictures and the moment of the choice which causes the higher detection rate in the EL condition.

(B): (2) is due to (b). It is the large number of questions which triggers detailed descriptions of the choice processes in the EL condition.

From (A) “The fact that participants are periodically asked to recall the pictures and the moment of the choice causes the higher detection rate in the EL condition”, we deduce (A’): “The fact that participants are periodically asked to recall the pictures and the moment of the choice in the EL condition *induces an act of recall* of this moment, which triggers the detection of the manipulation.”

The recall instructions still enabled the participants to detect the manipulation after 40 min of interview in 80% of cases (this much longer gap between the two presentations of the pictures might have led to a much poorer detection rate). This shows that the act of recall was still efficient after that time gap. We can therefore conclude (C) that the descriptions of the choice processes given in the course of the interview relied on reliable memories of the moment of the choice, which strongly bears out the validity of these descriptions.

From (B) “It is the large number of questions which triggers detailed descriptions of the choice processes in the EL condition”, we deduce (B’): “The large number of questions *induces an orientation of the subjects’ attention towards their choice processes* which causes detailed descriptions of these processes in the EL condition.”

We consider that these acts of recall and orientation of attention are the introspective acts which enable the participants to provide a description of their choice processes which is both detailed and reliable.

Our hypothesis is that in the NEL reports, the subjects do not perform any of the above acts. On the one hand, they do not evoke the moment of the choice. One of the reasons which may explain this absence of evocation is the social interaction in which they are engaged, which does not incite them to introspect. They simply provide a justification for their choice that makes use of current publicly available information and makes sense in the context of the current social interaction, which is based on an implicit relationship of trust in the interviewer. We hypothesize that the subjects content themselves with justifying their choice by commenting in real time the photograph they have in front of their eyes – which explains why Johansson et al. found no significant difference between the manipulated and non-manipulated reports. The implicit trust in the interviewer, as well as the question “why”, do not prompt them to retrieve precisely the moment of the choice, which might have led them to detect the manipulation. The memory has not faded, it is simply not summoned up. The viewing of the video recordings of the two types of reports supports this hypothesis: during the NEL reports, subjects look alternately at the interviewer and at the picture they have in front of them. On the contrary, during the EL reports, the subjects are encouraged to evoke the moment of the choice, which leads them to break off eye contact with the interviewer and look “into empty space” at the evoked pictures (in a location which may be different from the table, where the pictures were originally shown), this unfocusing of the eyes being one of the criteria of the evocation state. The loss of eye contact contributes to the creation of another type of social interaction, where the interviewer is very present through his prompts and questions, while at the same time authorizing the subject to “leave” the usual mode of interaction in order to come into contact with his internal processes – which enables him to detect the manipulation. In other words, in the NEL condition the (inauthentic) relationship of trust and the persistence of eye contact with the experimenter make it difficult for the participant to imagine that he could have been deceived, while the EL condition fosters the participant’s confidence in his own processes, which enables him to recognize the manipulation.

On the other hand, subjects in the NEL condition do not redirect their attention towards their choice processes. They keep focused on the characteristics of the faces they have in front of them, i.e. on criteria (of the choice they did not make). In the absence of any introspective skill or help of the interviewer, they are not aware that they have done anything particular in

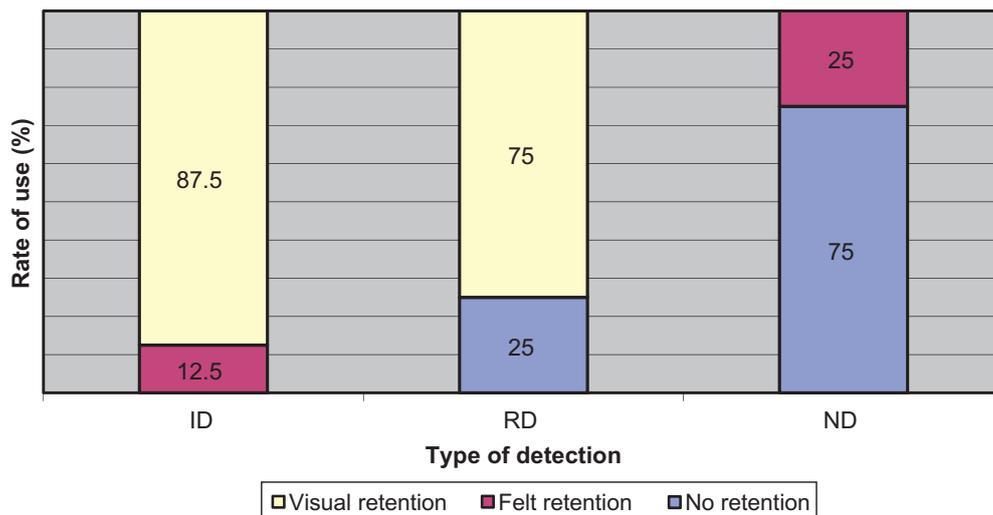


Fig. 4. Rate of use of retention per type of detection in the EL condition (ID = immediate detection, RD = retrospective detection, ND = no detection).

order to choose, nor that something particular may be done to retrieve it. Our hypothesis is that this absence of redirection, coupled to the absence of evocation, can explain the paucity of the reports, in addition to their unreliability.

Furthermore the absence of a “learning” effect – the fact that the detection rates in the NEL trials following an EL report do not increase – seems to confirm that it is not sufficient to carry out these acts of evocation and redirection once in order to know how to go about doing them, and that they are genuine skills which need to be learned.

It is true that a few NEL trials result in detections. However, the fact that these detections are preceded by especially long deliberation times (an average of 11 s while the general average time is 3.4 s) suggests that the very scarce detections in the NEL condition are due to a stronger memory trace due to more time spent examining the face at the moment of the choice, and not to an act of evocation of the faces during the report.

But how can we explain the fact that after the elicitation interview, some subjects detect the substitution and some others do not? Should the descriptions which did not lead to detection be considered as inauthentic? Among the four descriptions which did not lead to detection, three are particularly detailed, rich and above all *consistent*. This consistency, defined as the absence of contradiction and the complementarity and coherence of various elements of description within the report, is in our opinion a strong criterion of authenticity. As the elicitation interview is an iterative process, the report it results in is a jigsaw made up of scattered pieces of descriptions, which the analysis aims to piece together again. When this reconstitution results in a coherent sequence, it is hard to imagine that the initial description could have been constructed by the subject without the guiding thread of lived experience. This raised the question as to whether the lack of detection could be explained by the nature of the strategy of choice used. To try to answer this, we looked for possible correlations between the detection rate and the type of strategy. This revealed that all the cases of non-detection followed a “feeling-focused” strategy, whereas most immediate (7/8 i.e. 87.5%) and retrospective (9/11 i.e. 81.8%) detections followed a “features-focused” or mixed strategy. Moreover, all the subjects who instantly detected the substitution (8/8), and 75% of the subjects who detected it retrospectively (9/12), used retention at the moment of the choice. In 7 trials out of 8 immediate detections, and in all the cases of retrospective detection, this was a visual retention. However among the subjects who did not detect the substitution, only one used retention, and this was a felt retention (see Fig. 4).

The use of visual retention at the moment of the choice, which characterizes a features-focused or mixed strategy, therefore seems crucial in detecting the substitution. Its absence in the “feeling-focused” strategies might therefore explain why 3 trials out of 4 result in a non-detection. But the low detection rate after the NEL choices, which may be assumed to use the same proportion of visual retention as the EL choices, shows that retention is not sufficient in itself. Our hypothesis is that the process of evocation reactivates during the interview the retentions which were used at the time of the choice, and that this reactivation is necessary to retrieve the initial experience and thus to detect the substitution.

5. Conclusion

Our experiment confirms that naïve descriptions of our decision-making processes are usually poor and unreliable. However it shows that it is possible to access these processes by carrying out specific acts consisting in evoking the process and directing one’s attention towards its different dimensions. These acts make it possible to produce a description of the choice criteria and processes which is both rich and reliable, as evidenced by the high detection rate of the substitution in our experiments. The fact that we knew our reports were reliable meant that we were able to analyze their content and to detect regularities in the form of a set of choice strategies.

The experiments of Nisbett and Wilson, like those of Johansson et al. whose subjects were not using a truly introspective approach, cannot therefore be considered to have discredited in any way the possibility of experientially accessing our decision-making processes in a disciplined way and studying them rigorously and systematically. In this perspective, choice blindness can be considered as a particular case of our ordinary blindness to lived experience, which can be remedied using appropriate means or procedures (Depraz et al., 2003; Petitmengin, 2006, 2009b; Petitmengin & Bitbol, 2009). This possibility of disciplined description of our decision processes gives access to vast and almost unexplored deposits of data, and opens up several avenues of investigation. The first consists in deepening the “first person” study of decision making: repeating the same protocol with a larger number of subjects, designing protocols involving other types of choices. The second avenue consists in studying in more detail the acts enabling such a disciplined access: their nature, the means that make it possible to elicit them through an interview or possibly self-introspection, their subjective and objective (notably linguistic) criteria of realization, the evaluation of the level of expertise of the subjects and of the interviewers (Froese, Gould & Set, 2011), the adaptation of introspective techniques to the type of process under study and to the strategy being used. Such studies will lead us to better understand and control the processes which enable us to access our cognitive processes.

Acknowledgments

We would like to thank the reviewers of this article for their very useful comments. We also thank Céline Astruc, Michel Bitbol, Tom Froese and Anil Seth for their advice and remarks. We are also grateful to Dominique Fréard for his help in performing the statistical analyses.

Appendices A and B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.concog.2013.02.004>.

References

- Bernard, J. M. (2003). Analysis of local or asymmetric dependencies in contingency tables using the imprecise Dirichlet model. In *Third international symposium on imprecise probabilities and their applications (ISIPTA'03)*. Lugano, Switzerland, July 14–17.
- Binet, A. (1903/2004). *L'étude expérimentale de l'intelligence*. Paris: L'Harmattan
- Braboszcz, C. (2012). *Study of the electroencephalographic correlates of mindwandering and meditation*. PhD Thesis, Université Paul Sabatier, Toulouse.
- Cahour, B., Brassac, C., Vermersch, P., Pachoud, B., & Salembier, P. (2007). Etude de l'expérience du sujet pour l'évaluation de nouvelles technologies; l'exemple d'une communication médiée. *Revue d'anthropologie des connaissances*, 1(1), 85–120.
- Cahour, B., Forzy, J.F., Martin, C. (2010). Feelings and strategies of senior drivers: Ways of coping with fear? In *Proceedings of ECCE 2010 (European Conference on Cognitive Ergonomics)*. Delft, The Netherlands, August 24–27.
- Dennett, D. C. (2002). How could I be wrong? How wrong could I be? *Journal of Consciousness Studies*, 9(5–6), 13–16.
- Depraz, N., Varela, F., & Vermersch, P. (2003). *On becoming aware. The pragmatics of experiencing*. Amsterdam: John Benjamins.
- DSM IV (1994). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Press Inc.
- Elster, J. (1983). *Sour grapes: Studies in the subversion of rationality*. Cambridge University Press.
- Fenichel, O. (1946). *The psychoanalytic theory of neurosis*. London: Routledge.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Freud, S. (1991). *Case histories II*. London: Penguin Books.
- Froese, T., Gould, C., & Seth, A. K. (2011). Validating and calibrating first- and second-person methods in the science of consciousness. *Journal of Consciousness Studies*, 18(2), 40–64.
- Gusdorf, G. (1951). *Mémoire et personne*. Paris: PUF.
- Hendricks, M. (2009). Experiencing level: An instance of developing a variable from a first person process so it can be reliably measured and taught. *Journal of Consciousness Studies*, 16, 129–155.
- Husserl, E. (1925/2001). *Analyses concerning passive and active synthesis: Lectures on transcendental logic* (Boston, MA: Kluwer Academic Publisher).
- Jack, A., & Roepstorff, A. (2003). Why trust the subject? Part 1. *Journal of Consciousness Studies*, 10(9–10).
- Jack, A., & Roepstorff, A. (2004). Why trust the subject? Part 2. *Journal of Consciousness Studies*, 11(7–8).
- Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, 310(5745), 116–119.
- Johansson, P., Hall, L., Sikström, S., Tärning, B., & Lind, A. (2006). How something can be said about telling more than we can know: On choice blindness and introspection. *Consciousness and Cognition*, 15, 673–692.
- Katz, E. (2011). Attending to clinical practice: A phenomenological study exploring the structure of clinical attention and its relationship with holistic competence. PhD thesis, University of Toronto.
- Light, A. (2006). Adding method to meaning. A technique for exploring people's experience with technology. *Behaviour and Information Technology*, 25(2), 175–187.
- Maurel, M. (2009). The explicitation interview: Examples and applications. *Journal of Consciousness Studies*, 16(10–12), 58–89.
- Mele, A. (2001). *Self-deception unmasked*. Princeton: Princeton University Press.
- Mill, J. S. (1882/1961). *Auguste Comte and Positivism*. Ann Arbor, MI: University of Michigan.
- Mills, J. (1958). Changes in moral attitudes following temptation. *Journal of Personality*, 26, 517–531.
- Myers, D. G., & Bach, P. J. (1976). Group discussion effects on conflict behavior and self-justification. *Psychological Reports*, 38, 135–140.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Overgaard, M. (2006). Special issue on introspection. *Consciousness and Cognition*, 15(4).
- Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193–198.
- Petitmengin, C. (1999). The intuitive experience. *Journal of Consciousness Studies*, 6(2–3), 43–77.
- Petitmengin, C. (2001). *L'expérience intuitive*. Paris: L'Harmattan.
- Petitmengin, C. (2006). Describing one's subjective experience in the second person. An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5, 229–269.
- Petitmengin, C. (2009a). Ten years of viewing from within. *Journal of Consciousness Studies*, 16(10–12).

- Petitmengin, C. (2009b). Editorial introduction. *Journal of Consciousness Studies*, 16(10–12), 7–19.
- Petitmengin, C. (Ed). (2011). Ten years' viewing from within: Further debate. *Journal of Consciousness Studies*, 18(2).
- Petitmengin, C. (2011). Describing the experience of describing? The blind spot of introspection. *Journal of Consciousness Studies*, 18(1), 44–62.
- Petitmengin, C., & Bitbol, M. (2009). The validity of first-person descriptions as authenticity and coherence. *Journal of Consciousness Studies*, 16(10–12), 363–404.
- Petitmengin, C., Bitbol, M., Nissou, J. M., Pachoud, B., Curalucci, C., Cermolacce, M., et al (2009). Listening from within. *Journal of Consciousness Studies*, 16(10–12), 252–284.
- Petitmengin, C., Navarro, V., & Le Van Quyen, M. (2007). Anticipating seizure: Pre-reflective experience at the center of neuro-phenomenology. *Consciousness and Cognition*, 16, 746–764.
- Remillieux, A. (2010). *Explicitation des connaissances de conduite du changement à la SNCF: Vers une gestion des connaissances pré-réfléchies*. Editions Universitaires Européennes.
- Remillieux, A., Petitmengin, C., Ermine, J. L., & Blatter, C. (2010). Knowledge sharing in change management: A case study in the French railways company. *Journal of Knowledge Management Practice*, 11(3).
- Schwitzgebel, E. (2007). No unchallengeable epistemic authority, of any sort, regarding our own conscious experience – Contra Dennett? *Phenomenology and the Cognitive Science*, 6, 107–113.
- Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology*, 37, 228–243.
- Sully, J. (1881). Illusions of introspection. *Mind*, 6(21), 1–18.
- Titchener, E. B. (1910/1980). *A textbook of psychology*. New-York: Scholars'Facsimiles and Reprints.
- Valenzuela Moguillansky, 2012. *The relation between pain and body awareness: A, investigation using experimental and experiential methods*. PhD thesis, University Pierre et Marie Curie, Paris.
- Varela, F. J., & Shear, J. (Eds.). (1999). *The view from within. First-person approaches to the study of consciousness*. London: Imprint Academic.
- Vermersch, P. (1994/2011). *L'entretien d'explicitation*. ESF: Paris.
- Vermersch, P. (1999). Introspection as practice. *Journal of Consciousness Studies*, 6(2–3), 17–42.
- Vermersch, P. (2000). Conscience directe et conscience réfléchie. *Intellectica*, 31, 269–311.
- Vermersch, P. (2004a). Modèle de la mémoire chez Husserl. 1/Pourquoi Husserl s'intéresse-t-il tant au ressouvenir. *Expliciter*, 53, 1–14.
- Vermersch, P. (2004b). Modèle de la mémoire chez Husserl. 2/La rétention'. *Expliciter*, 54, 22–28.
- Vermersch, P. (2006). Rétention, passivité, visée à vide, intention éveillante. *Phénoménologie et pratique de l'explicitation*. *Expliciter*, 65, 14–28.
- Vermersch, P. (2009). Describing the practice of introspection. *Journal of Consciousness Studies*, 16(10–12), 20–57.
- Vermersch, P. (2012). *Explicitation et phénoménologie*. Paris: Presses Universitaires de France.
- Weisberg, J. (2011). Describing inner experience. *Journal of Consciousness Studies*, 18(1).
- Wilson, T. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Harvard University Press.