A Defense of Introspection from Within

Michel Bitbol • Archives Husserl (ENS), Paris, France • michel.bitbol/at/ens.fr
Claire Petitmengin • Institut Mines-Télécom (TEM) and Archives Husserl (ENS), Paris, France • claire.petitmengin/at/telecom-em.eu

> **Context** • We are presently witnessing a revival of introspective methods, which implicitly challenges an impressive list of in-principle objections that were addressed to introspection by various philosophers and by behaviorists. > **Problem** • How can one overcome those objections and provide introspection with a secure basis? > **Results** • A renewed definition of introspection as "enlargement of the field of attention and contact with re-enacted experience," rather than "looking-within," is formulated. This entails (i) an alternative status of introspective phenomena, which are no longer taken as revelations of some an sich slice of experience, but as full-fledged experiences; and (ii) an alternative view of the validity of first-person reports as "performative coherence" rather than correspondence. A preliminary empirical study of the self-assessed reliability of introspective data using the elicitation interview method is then carried out. It turns out that subjects make use of reproducible processual criteria in order to probe into the authenticity and completeness of their own introspective reports. > **Implications** • Introspective inquiry is likely to have enough resources to "take care of itself." > **Constructivist content** • It is argued that the failure of the introspectionist wave of the turn of the 19th/20th centuries is mostly due to its unconditional acceptance of the representationalist theory of knowledge, and that alternative non-representationalist criteria of validity give new credibility to introspective knowledge. > **Key words** • Introspection, first-person, phenomenology, consciousness, awareness, epistemology, non-representationalism, experience, subjectivity.

Introduction

Two extreme positions have been held about the reliability and accuracy of introspective research (Hurlburt & Schwitzgebel 2007). One position consists of claiming some sort of *infallibility* of direct access to one's own lived experience. It was adopted by several major philosophers, including Descartes (1990), John Locke (1975), David Hume (1978), Edmund Husserl (2004), and Alfred Ayer (1946), who thereby wished to secure an absolute foundation for knowledge. It was also defended by some contemporary functionalist or reductionist philosophers of mind with a negative undertone: according to them, some judgments about one's own experience are *trivially* true either because they are self-fulfilling (Chalmers 2003; Horgan & Kriegel 2007; Tye 2009), or because they are *expressive* rather than interpretative (Wittgenstein 1958; Bar-On 2004), or because there is no way to challenge them (Dennett 2002). Another, apparently opposite, position, consists of drawing up an overwhelming list of objections against any knowledge claim based on introspection. It is thus argued that introspective judgments are mere inferences rather than direct perceptive pieces of information (Wilson &

Dunn 2004), that they do not provide us with access to our own mental processes better than to those of others, and that they are biased and incomplete in view of the unconscious nature of a large part of mental processes (Merickle, Smilek & Eastwood 2001; Becker & Pashler 2002), incapable of disclosing the true motives of our acts, or even systematically self-deceptive in so far as they are little more than *ex post facto* confabulations (Nisbett & Wilson 1977). It is further pointed out that if introspection does not purport to make judgments, but only sticks to lived experiences, then it can hardly be taken as a piece of "knowledge" (Prinz 2004).

This catalogue of harsh criticisms tended to substantiate the belief that the disappearance of introspectionist psychology during the first quarter of the twentieth century (Lyons 1986; Danziger 1994; Costall 2006) was due to fundamental reasons that condemn introspection forever. The rise of cognitive sciences and neuroscientific approaches could then be taken as an additional reason to dismiss introspection as an old-fashioned and weak methodology in probing mental operations. But closer analysis shows that introspection was never really abandoned in psychology as a pri-

mary source of information or as a guide to behavioral and neuroscientific hypotheses (Boring 1953; Kriegel 2013), that some objections against it could be addressed successfully (Ericsson & Simon 1984), and that it is presently undergoing a phase of overt rebirth under various guises (Vermersch 1994; Depraz, Varela & Vermersch 2003; Hurlburt & Heavey 2006; Barrett et al. 2007; Marti et al. 2010). Far from being discarded by advances in neurosciences, introspection is increasingly seen as a full-status partner of such advances (Varela 1996; Rudrauf et al. 2003; Lachaux 2011).

The new wave of introspective studies prompts us to challenge traditional objections and to transform them into themes of empirical inquiry. For instance, is it true of *every* introspective report that it relies on indirect inference, rather than direct acquaintance? Is it the case that a vast majority of mental processes are irretrievably "unconscious"; or are many of them merely unattended and therefore not immediately apparent in explicit memories, yet available for being called forth if certain conditions are fulfilled? Can one not create introspective situations wherein any temptation to confabulate is prevented? Besides, could not one avoid the above-mentioned oscillation

269

270

between infallibility and systematic failure if an intermediate status is conferred on introspective material? The two standard statuses are either immediate *adhesion* to what is lived presently, or *judgments, beliefs, and rationalizations* about one's own (just) past cognitive processes. The intermediate status we favor in this paper is tantamount to probing iteratively into one's own *re-enacted* past experience, describing it at the lowest level of conceptualization, and then delegating the task of retrospective categorization to somebody else. The problem is that turning away from the usual alternative between being and knowing, between mere adhesion to what is lived and "judgment about something detached from the judging subject," requires a careful re-examination of our epistemological standards. One must renounce the classical requirement of "detachment" and "representational accuracy" without losing the benefit of objectifying (in the weak sense of looking for intersubjectively reproducible structures). Accordingly, in introspective studies, the usual "correspondence theory of truth" must be replaced to a certain extent with other, more immanent and more flexible criteria of validity, such as the criterion of "performative coherence." These epistemological issues will be reviewed in Part 1.

Now, as soon as the new criterion of performative coherence is adopted, it is tempting to implement and test it by means of empirical introspective studies. In Part 2, which represents the bulk of this paper, we thus present a preliminary study of the performative coherence of introspective reports, made by using the elicitation interview technique[1] (Vermersch 1994; Petitmengin 2006) on several subjects. The process of adjustment between re-enacted experiences and descriptions of interviewed subjects is carefully analyzed, and generic structures of this adjustment are extracted. This offers an internal probing of introspection: an assessment of introspection by its own critical resources.

1| This expression translates the French "entretien d'explicitation" used by Pierre Vermersch and Claire Petitmengin. It has recently been chosen as the "official" English name of the method. Indeed, the word "explicitation" is not used in English.

Part 1: In-principle objections and replies – Can introspection be impossible yet real?

In this section, we shall list and briefly discuss some of the main epistemological features that were missing (or poorly understood) in early introspective psychology, and that may explain its vulnerability to criticism (Bitbol & Petitmengin 2013a, 2013b). We also wish to display, whenever possible, the ability of the kind of introspective studies we practice to cope with most of these deficiencies. We will list four such objections (Petitmengin & Bitbol 2009; Vermersch 1999), and outline some replies new introspection has in store for them.

Objection 1

The most archetypal objection is that it is impossible to observe one's own experience, because this presupposes a split between subject and object, while in this case the object is nothing else than the subject itself. Yet this kind of objection is directed against introspection as prejudice says it *should be*, rather than against introspection as it *is in fact* practiced. The prejudice is that part of the subject engages in second-order observing or monitoring of first-order mental processes. But against this prejudice, several results are consistent with the idea that introspection merely involves a modified version of those very first-order mental processes. Indeed, it was found that giving subjects the instruction to introspect in addition to perceiving, modifies the early perceptual processes as revealed by the Event Related Potential technique, rather than merely altering post-perceptual processes involving memory (Overgaard et al. 2006). As we will see later on, this is also supported by phenomenology, where a replacement of the standard idea of self-observation by "non-observational awareness" (Marcel 2003) is advocated.

Objection 2

According to the second set of objections, introspection alters the mental process to be known. Two major varieties of alterations are: (i) observational, since the act of introspective examination *disturbs* the

mental flux; and (ii) interpretative, since the categories used for describing or explaining what is lived are theory-laden.

However, these can only be called "alterations" or "distortions" with respect to experience as it is in itself, prior to any attempt at observing, catching, and interpreting. In other words, the previous objections rely on some version of the myth of the "given" (Garfield 1989). Distancing ourselves from this myth, a very different picture arises. Speaking of a process *an sich* that is unfortunately disturbed by the coarse instrument we use in order to have access to it only makes sense if there is a way of accessing it independently of this coarse instrument (Jack & Roepstorff 2002). But if there is nothing to compare with the instrumental outcomes, this is speculation. True, one could try to use several types of introspective methods, subtract variations from one method to another, and thus hope to identify some stable (structural) feature that *plays the role* of undistorted experience. But since no *non-*introspective method can even in principle teach us what *is* our experience in itself, we have no other choice than to accept the latter stable feature as our only candidate for the title of "described experience." This strategy of substitution is similar to Bohr's approach of quantum mechanics. According to him, speaking of "disturbance of phenomena by observation" is confusing, and one should rather *define* the phenomenon as the indivisible byproduct of the observation process.

Besides, recent methods of verbal report and introspection are meant to take the disturbance objection into account. Thus, the elicitation method (Vermersch 1994; Depraz, Varela & Vermersch 2003; Petitmengin 2006; Petitmengin et al. 2009) that we currently practice can be characterized as a strategy for progressively unfolding initially "pre-reflective" aspects of lived experience by asking subjects to rehearse and even to *re-enact* this experience while broadening their field of attention. Here, *retrospection* (as opposed to "thinking-aloud" protocols) is systematically used. This addresses the traditional objection that observation disturbs the observed process if it occurs simultaneously to it. Moreover, this enables patient expansion of awareness, part after part, of a selected slice of experience. As for the issue of the interpretative distortion, it

is pragmatically addressed (see next point) by asking subjects to lower as much as possible the level of conceptualization of their descriptions.

Objection 3

The third set of objections amounts to claiming that one is systematically mistaken about one's own experience. Part of this objection is grounded on the observation that it is very easy for subjects to go astray about the *stimulus* that was presented to them in order to trigger a certain experience (Titchener 1912; Schwitzgebel 2004). Criticisms have been formulated against the propensity subjects have towards saying that they see more than they can evidence (Dennett 1992), or against their inability to see major parts of what occurs in front of them if their attention is distracted (as shown by experiments on "change blindness," Silverman & Mack 2006). However, this charge might well be excessive or misplaced. In a non-representationalist epistemological framework, the issue of the truth or reliability of introspective descriptions is likely to be given a completely new meaning. This was already suggested in a pioneering paper by Benny Shanon (1984), and in an increasing number of articles since then (Piccinini 2003; Piccinini 2009; Petitmengin & Bitbol 2009). These authors pointed out that standard critiques just show that introspective data cannot always be evaluated on the basis of correspondence. This is not to be regretted, since, after all, no other data, including in experimental science, are in fact evaluated this way either. Experimental data are accepted as reliable not when they are seen to correspond with reality "in itself," but as soon as they become…

" effectively frozen at moments of interactive stabilization … encompassing, on the one side, captures and framings of material agency, and, on the other, regularized, routinized, standardized, disciplined human practices. " (Pickering 1995: 102)

The alternative is then evaluation on the basis of performative coherence, where "coherence" can concern several levels of practice: internal coherence in self-assessment and report; interpersonal coherence in dialogue; and triangulated coherence in a network connecting introspective reports with experimental practice as prescribed by neurophenomenology (Varela 1996). Just as, according to the second Wittgenstein, language must take care of itself (without foundational security in logic), introspection must take care of itself as much as possible (without foundational security in "correspondence" of any kind). The major part of this article (part 2) is aimed at defining some tentative criteria of internal and interpersonal coherence, thus offering a basis of validation adapted to the case of introspection.

This retreat from the correspondence theory of truth to an extended version of the coherence theory of truth, however, should not be overrated. Firstly, it does not exclude the use of the word "correspondence" as a convenient summary of the "choreography" (Pickering 1995) of acts that contribute to maximizing performative coherence. The elusive quest for "correspondence" can even be considered as a good motivation, or "regulative ideal," for reaching an optimal form of coherence, and thereby a piece of reliable knowledge. In other words, the concept of correspondence can be considered as part and parcel of the quest for coherence. Secondly, as a consequence, it is perfectly licit to look for types of performative coherence that invite interpretation as signs of improved correspondence between an introspective report and the experience it is meant to describe. It has thus been shown that one can considerably improve the faithfulness of first-person reports in the precise experimental situation that has been taken for more than thirty years as the archetypal rebuttal of introspection, namely in the Richard Nisbett & Timothy Wilson (1977) setting as adapted by Petter Johansson et al. (2006). This improvement, which raises the agreement between an initial experience of choice of presented faces and the later report of this experience from about 30% to about 80%, has been obtained by inserting an elicitation interview between the moment of the choice and the moment of the final report (Petitmengin et al. 2013).

Objection 4

The fourth and final group of objections focuses on the purely subjective status of introspective descriptions, and on the fact that the situation it concerns is irreproducible. If this is so, a verbal report of introspection only concerns the person that reports at a certain time; it teaches us nothing about other persons, and perhaps not even about oneself at any other time. This is a serious objection, but a renewed conception of objectivity that arises from a non-representational view of science suffices to meet it.

The sought alternative definition of objectivity is borrowed from Kant. In it, objectivity is not something to be found ready-made *out there*, but a project of operational extraction of invariant structures from a cluster of appearances. So, the issue as to whether or not single events teach us something objective is to be decided on a methodological, *not* on a metaphysical plane. What method should we choose in order to reach objectivity *qua* invariance?

In any science, extracting invariant or covariant structures relies on a process of ascent in generalization and theoretical abstraction, symmetrical to the process of descent that is necessary to reach a nucleus of discourse that can be considered as "factual" or "data-like." The procedure should be the same for introspection. It should involve:

- Descent towards minimally interpreted descriptions of the subtlest lived events, without any attempt at asking the *subject* to reconstitute her own cognitive processes, or to explain her "reasons" *in abstracto*. This is tantamount to inducing a very careful process of phenomenological reduction in the introspecting subject.
- *A posteriori* ascent of the *scientists* that are analyzing a corpus of introspective reports construed as data towards structures generic enough to be seen as stable and invariant across subjects and circumstances. As Shanon (1984) cogently pointed out, "structures are less particular than content: they are not associated with the introspector's idiosyncratic experiences, nor are they likely to be affected by the process of data collection itself."

This two-step procedure is exactly the one we apply when we practice the method of elicitation of experience by interviews:

1 | being very careful in guiding subjects towards exquisite contact with their experience and undoing any rational reconstructions or generalizations that

may interfere with their task of description;

2 | retrieving the data extracted from many such disciplined descriptions and extracting generic structures out of them (Shanon 2003).

We gather from these objections and sketchy replies that the most crucial weakness of the introspectionist wave of the turn of the nineteenth and twentieth centuries is likely to be its unconditional acceptance of the classical, representationalist, theory of knowledge, despite many criticisms stemming from philosophy of science (Pickering 1995) and cognitive science (Varela, Thompson & Rosch 1991; Thompson 2007). In the following section, we develop a preliminary internal defense of introspection that deliberately avoids the pitfalls of representationalism.

Part 2: Introspection of introspection – How first-person reports can take care of themselves

We wish here to present a preliminary piece of empirical work in assisted introspection that shows how the criterion of "internal coherence" can be used as some sort of warrant for reports of first-person experience.

Why a warrant for first-person experience is needed

To start with, it is useful to come back to a remark made in the introduction. Some thinkers from the phenomenological tradition may perceive the very idea of seeking a warrant for first-person experience as strange or unnecessary, especially if they overlook the differences between phenomenology and introspection. After all, they may notice, phenomenology traces back the very *origin* of any warrant to pure experience. What better warrant do we want than pure experience itself? This claim is apparently supported by Husserl's statement of a phenomenological gap between physical objects and the field of lived experience (Husserl 2004):

- Physical objects are given by way of (their) angles and profiles, which are

associated with expectations about aspects that could be apprehended in the future. Husserl calls the set of expectations about future aspects of an object a "horizon." This means that physical objects are always *incompletely* given, and that they are therefore "essentially" subject to doubt. Indeed, there is a possibility that future contents of experience disconfirm previous conjectures about these objects.

- By contrast, writes Husserl, lived experience is *completely* given. Future experience cannot *disconfirm* its existence here and now, but only replace it. Lived experience is then an "absolute," in so far as its existence is certain and indubitable while it is present.

The latter proposition (which has often been challenged in view of its alleged "presentism" and "foundationalism") is striking, and it appears relevant for any approach to first-person experience. But it must be discussed in order to elucidate its exact meaning. One should first understand that the kind of certainty or indisputability that is inherent to lived experience has nothing to do with an ordinary certainty *about* something. It is a certainty without aboutness, a certainty that coincides with its content; it is a blank, unformulated, and unintentional certainty. By contrast, as soon as one sets out to explore reflectively that *about which* there is a certainty, a new experience arises; a distinction is then established between this new experience and the domain of experience it purports to explore. In this latter case, there is no more "absoluteness" or "inherent warrant" in reflected experience than in anything else.

The reason for this absence of warrant for reflected experience can be described more precisely by way of an analogy with ordinary perception. Reflection on one's own experience may bring out "margins" or "fringes" (James 1890) of the field of attention. These margins are initially almost implicit (in so far as we are pre-reflectively unaware of them) but they progressively unfold in the course of a reflective exploration. While the attention is focused, and these margins are still confined to the implicit, one can nevertheless have *expectations* about their future reflective unfolding, so much so that at every given moment

they are almost coextensive to these expectations (Mangan 2007). In the same way, in ordinary perception, the presently visible aspect of a physical object is surrounded by a "horizon" of expectations about future unfolding aspects or profiles.

This being granted, two questions arise, and two quick answers can be given:

1 | Can we disconfirm the expectations we have about these pre-reflective margins of experience, just as we can disconfirm our expectations about future profiles of physical objects? The answer to this question is plainly "yes," because we all know cases where we do *not* find what we expect of ourselves[2] in the course of an accurate reflective exploration.

2 | Is it correct to say that, when we undertake a reflective move about our own experience, we only "*reveal*" its "sides and margins" that just happened to escape the light of awareness? The answer to this question should be nuanced. Indeed, on the one hand what we are tempted to take as a mere "revelation" of hidden aspects of experience arises in fact as a new experience. On the other hand, this new experience is no arbitrary mental production; it comes with the internal characteristic of being a re-enactment and unfolding of some previously lived experience, and the accuracy of this claim can be checked experimentally (Petitmengin et al. 2013).

But if disconfirmation (in the sense of a discrepancy between initial expectations and later reflective experiences) is possible, and if there is no true "revelation" of the margins of experience, the original, coincident, "certainty" documented by Husserl is no longer with us. Warrant *has* to be sought for reflection, even though it can be found nowhere else than in experience itself. More precisely, warrant for reflected contents can be found nowhere else than in an experience of *mutual consistency* between the steps and outcomes of the reflective process. So, let us now expound how this experience of internal coherence can be elicited.

2| The practice of psychoanalysis, and our practice of elicitation interviews (see below for a description), tend to substantiate this elementary finding of fallibility about oneself.

Elicitation: An interview method of assisted introspection

Our research about the reliability of first-person reports used an interview method that was explained at length in a previous article (Petitmengin 2006). It was especially designed (Vermersch 1994) to overcome certain pitfalls of turn-of-the-century introspection, and of many other attempts since then. The interviewing process (whose duration ranges between 20 minutes and one hour) essentially consists of leading the subject to realize the following acts:

▪ Moving from a general representation to a singular experience (usually to a very thin time-slice of this experience). Indeed, what is to be reported, made explicit, and carefully reflected upon, is a past singular *lived* experience, *not* "the experience I *believe* I *generally* have when this *kind* of situation occurs." This option, implemented by asking the subject to think about a particular event in her life and helping her to stick to it, enables true intimacy with the experience to be described.

▪ Repeatedly *re-enacting* and thereby *evoking* selected time-slices of this singular experience, using to that purpose concrete episodic memory, rather than abstract memory. Guiding the subject towards these acts relies on simple methods such as helping him/her to retrieve the spatio-temporal context of the experience and the sensations associated with the experience, or systematically using the present tense (which is also one of the many verbal or bodily signs we use to check that the subjects are indeed in the dimension of re-enactment during the interview).

▪ *Stabilizing* attention on the singular experience by:

1 | inviting the subject to suspend any other concern than the present interview

2 | reformulating the last pieces of report given by the subject (as an answer to a previous question) and inviting her to check the accuracy of the reformulation

3 | bringing the subject back to the thread of her description each time she stops describing the experience and starts judging or explaining it

4 | isolating a relevant feature of experience from its background, through a word or a gesture that singles it out.

▪ Relaxing any attention focused on the "why" and "what" of experience ("Why did you feel so?"; "What were you aiming at then?"), and progressively expanding it to appreciate the "how" ("How exactly did you carry out this task?" "How did you know?"). In other words, moving attention from the narrow *content* to the complete *act* of consciousness, which is tantamount to performing the phenomenological reduction. This broad contact-like form of attention is induced by patiently asking the subject many how-questions about every aspect of the experience, and pulling her back from any tendency to abstraction or generalization each time this manifests itself.

▪ Directing attention to the various dimensions and aspects of the selected singular experience, one after another, during successive runs of its re-enactment. For example, by encouraging the subject to describe successively the visual, auditory, kinesthetic, and emotional features of the experience to be reported, including details that were not noticed at the moment when that experience was lived for the first time. The interviewer can be helped in identifying which sensorial dimension the interviewee is presently exploring by carefully observing her eye movements (Buckner, Reese & Reese 1997). She can also help the interviewee by asking many neutral and virtually contentless questions aimed only at changing slightly the direction of attention.

Interviews about the accuracy of interviews

The data we are to discuss below are extracted from a series of 13 interviews about the experience of listening to a sound (a brief natural or electronic sound, or a short piece of music). These interviews are scrutinized in detail in another article (Petitmengin et al. 2009). But here, we wish to focus on two meta-questions that were asked at the very end of each one of the interviews, and that aimed at evaluating the felt *accuracy* of the resultant introspective reports. The answers to those meta-questions were not accounted for in the initial article, but we shall study

them here in view of their special epistemological relevance. The questions are:

a   Do you have the feeling that the report you have given until now conforms to the experience you were describing?

a′ Do you have the feeling that the report you have given until now is complete?

b   How do you know that your report does (does not) conform to the experience?

b′ How do you know that your report is (is not) now complete?

Once again, questions about conformity of reports to experience by no means imply, of course, that we adhere to the correspondence theory of truth. They do not imply that we naively ask subjects to compare their description to their *initial* experience. They are only to be taken as performatives that invite the subjects to do what they can to formulate an appropriate answer: exploring their combined *present* experiences of re-enaction and description.

But another methodological issue must be clarified before we come to the results. After the descent to close contact with the very flesh of their experience (and away from abstraction) that is prescribed *to subjects*, an ascent towards general concepts and structures, and therefore towards objectivity (as redefined in Part 1), is undertaken *by the experimenters*. As soon as the dialogues that contain elements of reports of singular experiences have been recorded and transcribed, the material is reorganized, resolved, and somehow formalized in order to display the subjective and intersubjective structure of experience. This material is rich and dense, the style and the particular expressions used can be highly personal, but it is possible to extract from there a meaningful structural residue. An accurate method to reach this aim by segmental analysis of verbal reports was described in previous work (Petitmengin 2001, 2006), and we will make use of it in what follows. However, due to the restriction we intentionally put on our material (that only concerns the answers to *a few* questions of epistemological import), we have reformulated and simplified the preliminary steps of this method by due reference to Claude Levi-Strauss's analysis of myth (Levi-Strauss 1958).

The discourse of myth, says Levi-Strauss, is characterized by extreme translatability. What is important is only its meaning as a *story*. Neither its style nor the many external

273

Questions / Answers	Does your description *conform* to experience?	Is your description *complete*, exhaustive?
Yes	61 % (8)	
Yes, but: Some precisions or nuances are lacking, the description is too coarse.	39% (5)	17% (1)
No	0	83% (5)
Total	100% (13)	100% (6)

Table 1: Answers to the questions about the adequacy of the description.

274

circumstances that surround the nucleus of that story are truly relevant. The story itself can be reduced to very simple propositions that express elementary relations between beings and acts. Just as the basic constituents of language are called *phonemes* (smallest differentiable *sounds*), *morphemes* (smallest units with meaning: *words*), and *semantemes* (smallest units liable to be true or false: propositions), the basic constituents of a myth are called *mythemes* (smallest form of a proposition able to express a relevant relation between beings and acts). When these mythemes have been identified, they can easily be collected into categories, and these categories in turn organized into a table or a schema that displays a higher-order pattern of relations and brings us to the functional meaning of the myth.

We have used virtually the same procedure for the reports offered by our subjects during the interviews:

1 | Progressive simplification of pieces of discourse (which are often hesitating, long, and repetitive, which testifies to good re-enactment of experience) until we have reached what we propose to call a "*descripteme*"

2 | Categorization of *descriptemes* coming first from one subject, and then from several subjects

3 | Regrouping of the categories of *descriptemes* into tables showing their similarities, differences, or hierarchical relations.

Let us give a brief example, taken from one of our interviews. The original fragment of interview runs thus:

❝*H:* Do you have the feeling that your report conforms to your experience?

J: Er… I think so, yes… Yes, yes…

H: How do you know that?

J: Because it is *present*. I am only missing the sound itself, if I dare say. It's present. There is no interval left for anything else… You see, there is nothing else. There is what I said, and maybe this should be dug into deeper. But there is no space for anything else in that description. It is exactly that. I was immersed in it and I was even surprised to find things I did not expect; for instance those double pulsating movements, the small notes, and the general pulsation of music with a different tonality. Yes, yes… I could never have invented this…❞

Step (1) The *descriptemes* we extract from J's answers are the following:
- "Report conforms to experience"
- "Experience is present"
- "No interval left for anything else"
- "Surprised to find things I did not expect"

Step (2) The main categories corresponding to the descriptemes of this subject are:
- *Conformity to experience*
- *Effective re-enactment of experience*
- *Impossibility to add something* (without betraying experience); more generally, *fulfillment of the intention to meaning*
- Ability of the descriptions to bring to attention unexpected features of experience.

Now, to document **Step (3)**, we have to organize the categories according to their relations, and write down the complete tables of categories as they have been extracted from interviews. This will be done in the next section.

Generic structures of the self-assessment of first-person descriptions

Table 1 displays the answers to the initial question about adequacy of description.

This table brings out that no subject expressed no confidence in his/her description, but that a large proportion of subjects felt partial dissatisfaction with the precision of their reports. However, the quasi unanimous statement of *incompleteness* of the reports, expressed by those subjects that were asked, shows that most of them felt a gap between their re-enacted experience and the lack of exhaustivity of their reports, even when they had no or little suspicion about their reasonable accuracy. Incompleteness or coarseness but satisfactory accuracy is the general judgment subjects make about their own reports.

The following tables are more subtle because they extract categories out of several long and nuanced answers to the question "how?" Yet, surprisingly, they proved quite easy to establish because the how-categories turned out to be well-defined and repeatable across subjects. According to the answer provided by the subject to the previous question (about the conformity of the description), the "how" questions were the following: "How do you know that your description conforms to your experience?" or "How do you know that your description does not conform to your experience?" These questions elicited the verbalization (meta-description) of the practical *criteria* that enabled the subjects to evaluate the accuracy of their descriptions. Interestingly, these criteria did not concern the *content* of the description, but the description as a *process*. For example, the spontaneity of emergence of the evoked experience, the incremental nature of the fulfillment of evocation, the feeling of not being able to add anything to the verbal description, are clearly internal processual criteria, enabling the subjects to assess the accurate realization of mental acts. We identified two types of processual criteria, intended to assess the accuracy of two processes: on the one hand the process of *evocation or re-enactment* of the past experience; on the other hand the process of *verbal description* of the experience. The verbalization of these evaluation crite-

ria was often followed by the spontaneous description of the *acts* they were intended to evaluate (what the subjects did to evoke their experience, what they did to describe it verbally) or to trigger (what they did to adjust their evocation or description when the criteria were not satisfied). However, when these acts – which usually belong to a deeply pre-reflective dimension of experience – remained implicit and were not described spontaneously, we relied on the description of the verbalized criteria to elicit their description, through questions such as: "What did you do to make the evocation of your experience more intense?" or "What did you do to adjust your verbal description?"

To sum up, our question "How do you know that your description conforms to your experience?" did not trigger the comparison of two contents of experience (the content of the actual description and the content of the initial experience), but rather the evaluation of the actual acts involved in the description process. The processual criteria and acts concerning the evocation process are presented in Table 2. The processual criteria and acts concerning the verbal description process are presented in Table 3.

The first set of processual *criteria* that support the subject's conviction about the conformity of their description with experience (first column of Table 2) is related to the process of re-enactment of experience. It can be developed as follows:

a What we called the "spontaneity of the emergence of constitutive elements of the re-enacted experience" bears the mark of the involuntary. These elements pop up unexpectedly, unpredictably, beyond the control of the subjects, during the successive runs of the operation of re-enactment of experience triggered by the interviews. This absence of control is typical of the process of re-enacting experience by using concrete episodic memory, where the memory cannot be deliberately set off, but its spontaneous emergence is prepared by the rediscovery of sensations linked to the initial experience. This involuntary character clearly differentiates the process of evocation from that of trying through effort to remember something by using abstract memory.

Processual criteria		Acts assessed by the criteria in the first column	
a	*Spontaneity* of the emergence of elements of the evoked experience **(1)**	a'	Adopting a specific attentional disposition **(1)**
b	*Incremental* nature of the fulfillment of evocation of experience **(3)**	b'	Iterative coming into contact with experience **(2)**
c	"*Recognition*" of the elements brought out by the re-enactment. Feeling of *obviousness* **(2)**	c'	Connecting the two former acts **(1)**

Table 2: Criteria and acts of the evocation process.

Processual criteria		Acts assessed by the criteria in the first column	
Feelings of satisfaction			
a	*Feeling of not betraying* the experience by its description **(1)**	a'	Coming into contact with the experience **(1)**
b	*Feeling of not to be able to add anything* (without betraying the experience, or inventing something) **(3)**	b'	Trying to complete the description **(1)**
c	*Feeling that experience can be shared* with the interviewer through words **(1)**	c'	Comparing the evoked experience with the experience that is elicited in the interviewer by way of the description **(2)**
Feelings of dissatisfaction			
d	Feeling that words are too coarse, too caricatural. Feeling that the description is not sufficiently fine-grained **(6)**		
e	Feeling of amplification, enhancement of the experience **(3)**	e'	Using words to unfold and intensify experience **(1)**

Table 3: Criteria and acts of the verbal description process.

b What we called the "incremental nature of the fulfillment of evocation of experience" refers to the capacity of the subjects to enrich progressively their exploration and their description of experience, each re-enactment calling forth new elements and generating new descriptemes.

c "Recognition" is the peculiar mode of manifestation of the elements of an initially pre-reflective experience during the process of its re-enactment. Subjects have the feeling of becoming aware of very familiar features, already existent but that were not recognized as such before the moment of evocation elicited by the interview. "I pay attention to features that were buried, not really present, not manifest. But actually, this is the experience of music I *always* live (…). This conforms to what I am sure I have lived, but I was not aware I was living at the moment in which I was living it" (*JA*). This recognition is often felt as highly surprising. It goes along with a feeling of obviousness, of necessity, which very much differs from the faint feeling of "believing to remember something." It gives the re-enacted experience a very special quality of intensity, density, and depth, which is a central mark of pre-reflective experience when it emerges in the domain of awareness.

We must point out once again that the three former processual criteria do not involve anything like an external comparison between the re-enacted experience and the past experience. They are definitely internal to the process of re-enactment and to the mode of manifestation of the evoked items. These criteria are, to repeat, used to assess specific mental *acts*. In the process of re-enacting experience, subjects continuously and pre-reflectively carry out micro-operations and micro-adjustments in order to satisfy these criteria.

a' The first processual criterion of the first column of Table 2 – spontaneity of the emergence of the elements of the re-enacted experience – is used to assess the act of instituting a special type of attentional disposition that consists of opening up the attentional focus and relaxing any tension towards narrowly defined contents. The elements of the evoked experience cannot be aimed at directly, their emergence is a process that cannot be forced, but only elicited through the adoption of this open and receptive attentional disposition.

b' The second criterion – the incremental nature of the fulfillment of evocation – testifies to the accomplishment of another act: the iterative act of *coming into contact* that enables the *unfolding* of the experience to be described. Indeed, the act of recalling something is not realized once and for all: it is constantly repeated, and intensified. Trained subjects are well aware of certain experiential criteria allowing them to assess the accomplishment of this act, such as the feeling of an increased intensity of evocation and improved contact with their experience. When this contact fades away, trained subjects carry out micro-operations allowing them to "refresh" or intensify this contact; and interviewers know some signs that allow them to assess this contact (Petitmengin 2006) as well as procedures aimed at bringing their subject back to their experience. Usually, this "coming back" to experience gives rise to new elements of description. The incremental nature of the fulfillment of evocation is therefore a sign of the iterative process of re-enacting and contacting experience that underpins this fulfillment.

c' The third criterion – feeling of recognition – expresses the accomplishment of an act of connection between the two former acts: opening up attention and letting experience unfold in this context of open attention. It is within the field of an attention open to every aspect of experience that its unfolding may be seen as successful beyond any expectation that could have been made in the ordinary, narrow, state of attention.

The descriptions of these acts, as given by subjects, provide us with further confirmation that what is at stake in the validation of a description is not its "correspondence" with the past experience, but rather its *authenticity* rooted in the fitting accomplishment of a process made of a sequence of mental acts.

The second set of processual criteria that enable the subjects to assess the accuracy of their description (first column of Table 3) is related to the process of verbal description. Here, the answers are quite dispersed (perhaps because this class of statement was not oriented by a specific question), but they fall squarely into two subsets that are related to the ability or inability of words to describe or call forth experience.

The first subset of criteria, which are feelings of satisfaction called forth by the process of description, are either personal feelings of fulfillment of a meaning-intention, or communicational criteria. Personal feelings of fulfillment of a meaning-intention are of two types: (a) varieties of personal feelings of not betraying the experience by its description, and (b) varieties of personal feelings of not being able to add anything without betraying the experience or inventing something. As for communicational criteria, they manifest through (c), the feeling that an experience can be shared by means of words. Indeed, the communication power of descriptions exceeds their narrow semantic content in any concrete interview between two human beings that become attuned to each other: words have the power to arouse an experience in an interlocutor, and this experience in turn enables the interviewer and the interviewee to reach a new form of lived agreement. Here is a testimony about this latter process in one interview:

> We almost managed to agree, not so much about the precise definition of what we were saying, but by using words that were not appropriate but enabled us to understand what we had … not said. The words, somehow, were nothing more than the indication of … something else. (*J*)

The second subset of criteria consists of feelings of dissatisfaction or over-satisfaction with verbal descriptions:

d Varieties of personal feelings of being dissatisfied with the ability of verbal descriptions to do entirely justice to experience because words are too coarse, too caricatural, and description insufficiently fined-grained: "Experience is much richer, much more intense. For instance, I feel that bodily sensations were truly very present and more complex than what I said." (D)

e Varieties of personal feelings of being (over-)satisfied with the inherent dy-

namical power of verbal descriptions, even beyond the experience they are meant to convey. Some subjects have the feeling that the process of description has the effect of intensifying, of amplifying experience in a way: "This experience is so light, so subtle that, in order to put words on it I had to amplify it a little bit." (C.)

These criteria of partial satisfaction or dissatisfaction of subjects with the outcome of their own process of describing an experience are used to assess or trigger specific mental acts, of which we have collected a few descriptions.

a′ For example, the feeling of not betraying the experience by its description is the result of an act, which is repeated throughout the process of description, consisting of coming into contact with the experience to be described, through the process of evocation. When this contact is established, the accurate words come spontaneously, in a completely involuntary process. When this contact is lost, words lose their anchoring in experience and become empty. "If I am well-permeated by the experience, I can speak until the moment when, poof!, one is only in words, it no longer means anything, and I then come back to experience." (N.) It is this making contact with the non-verbal dimension of experience that ensures the accuracy of the description.

c′ The feeling that experience can be shared with the interviewer through words is the criterion of realization of another act, which consists of evaluating the capacity of words to call forth an experience in the interviewer. Several subjects described the act of figuring out rapidly and repeatedly which experience has been triggered in the interviewer by their description, and whether this triggered version fits with its re-enacted version in themselves. The latter act clearly shows that words are not evaluated with respect to their static capacity to *correspond* to an experience, but with respect to their dynamic capacity to *arouse* an experience, a process that supposes in turn the realization of specific acts by the person to whom the words are addressed.

e′ The feeling of amplification of experience is also the result of an act, which consists of using words in order to promote the unfolding of experience and enrich it with new nuances. Even though, as mentioned in point (d), words may be too coarse to do justice entirely to experience, they have the ability to help subjects to discriminate and intensify slight differences in experience, on which they can rely to refine their awareness of this experience. Words then play the role of "handles" that contribute to stabilizing a step of the re-enactment of experience and to giving easier access to further "incremental" steps of its unfolding.

Despite their coarseness, which prevents them from corresponding exactly to experience, words then have the power to amplify, discriminate, and promote unfolding of experience, as well as inducing an experience in the listener. It is this entire process that has to be evaluated to check the authenticity of the verbal description.

We infer from Tables 2 and 3 that an introspective description triggered by elicitation interview is simultaneously less and more than an expression of some original lived experience. It is a process in its own right that leaves aside certain features of experience by its coarseness yet helps disclose unsuspected features of experiences, marks their contrasts, and has potentialities for iterative self-corrections. Moreover, its value is not only cognitive and first-personal; it is strongly second-personal, since it enables one to convey (by verbal and non-verbal communication) many more aspects of lived experience than initially suspected on the basis of pure semantic analysis, and to increase the efficiency of this transmission by the stimulating effect of dialogue.

To recapitulate, the assessment of a description of experience obtained through the elicitation interview method does not rely on its comparison with past experience, but on internal criteria of accomplishment of mental micro-acts to which the subject has a present access. This gives rise to a statement of *authenticity* rather than "correspondence truth" of the description. The more a subject becomes an expert in re-enactment and description of experience, the more these acts and criteria are refined. Actually, it is future experience rather than

past experience that supports a description. Indeed, the process of re-enactment and unfolding allows the subject to identify pre-reflective aspects of her experience. Future experiences will corroborate this identification if they allow the subject to realize immediately, in real time, that such aspects are indeed present.

Conclusion

This preliminary study does not claim to be exhaustive, but to provide an outline of the performative criteria of validity of introspection, and to specify a line of research. It has two lessons in store for the future of introspection. The first lesson is that one must not neglect the information drawn from introspecting the introspective process. This information is precious material enabling one to feed back corrections on the methodology of interviews or individual phenomenological research, and to progressively elaborate reproducible and universally applicable procedures (Vermersch 1999; Depraz, Varela & Vermersch 2003). Without it, introspection is blind (Petitmengin 2011), and its practice can easily split into divergent schools. The second lesson is that introspective inquiry has enough resources to "take care of itself," namely to search for an optimal adjustment process between re-enacted experience and descriptions, thereby formulating an internal criterion of reliability and even "validity" (Petitmengin & Bitbol 2009). This procedure can certainly be supplemented by a quest for coherence on a broader basis (including neurological signatures); but it can by no means be *replaced* by it since, as neurophenomenological research has shown, neurological signatures can only be defined and disambiguated by interacting with careful introspective inquiry.

The past diffidence of psychology toward introspection is then likely to disappear in the same way as diffidence toward telescope images (an "artificial" lens-distorted product of vision, according to the Aristotelian natural philosophers) disappeared in astronomy during the seventeenth century: by the widespread use of this instrument, by the coherence and fruitfulness of the findings it allows, and by its efficient interconnection with other areas of knowledge.

277

MICHEL BITBOL

is presently "Directeur de recherche" at the Centre National de la Recherche Scientifique, Archives Husserl, Ecole Normale Supérieure, in Paris (France). He teaches philosophy of modern physics and epistemology. He received, successively, his M.D. in 1980, his Ph.D. in physics in 1985, and his "habilitation" in philosophy in 1997 in Paris. He worked as a research scientist in biophysics from 1978 to 1990. From 1990 onwards, he turned to the philosophy of physics, in which he edited texts by Erwin Schrödinger and developed a neo-Kantian philosophy of quantum mechanics. He also studied the relations between the philosophy of physics and the philosophy of mind, working in close collaboration with Francisco Varela. He is recipient of an award from the "Académie des sciences morales et politiques."

CLAIRE PETITMENGIN

After studies in Buddhist philosophy and ten years of consulting and research on information system design, Claire Petitmengin completed her Ph.D. thesis under the supervision of Francisco Varela at the Ecole Polytechnique in Paris on the subject of the lived experience that accompanies the emergence of an intuition. She is presently Professor at the Institut Mines-Télécom (TEM) and associate researcher at the Archives Husserl, Ecole Normale Supérieure, Paris. Her research focuses on the usually unrecognized dynamics of lived experience and "first-person" methods enabling us to become aware of it and describe it. She studies the epistemological conditions of these methods, notably the conditions of validity of their results, as well as their educational, therapeutic, artistic, and technological applications. Her research also addresses the process of mutual enrichment of "first-person" and "third-person" analyses in the context of neurophenomenological projects.

References

Ayer A. J. (1946) Language, truth, and logic. Gollancz, London. Originally published in 1936.

Bar-On D. (2004) Speaking my mind. Oxford University Press, Oxford.

Barrett L. F., Mesquita B., Ochsner K. N. & Gross J. J. (2007) The experience of emotion. Annual Review of Psychology 58: 373–403.

Becker M. & Pashler H. (2002) Volatile visual representations: Failing to detect changes in recently processed information. Psychonomic Bulletin & Review 9: 744–750.

Bitbol M. & Petitmengin C. (2013a) The science of mind as it could have been: About the contingency of the (quasi-)disappearance of introspection in psychology. In: Soler L. & Trizio E. (eds.) Science as it could have been. Routledge, London. In press.

Bitbol M. & Petitmengin C. (2013b) On the possibility and reality of introspection. Kairos 6: 173–198.

Boring E. G. (1953) A history of introspection. Psychological Bulletin 50: 169–189.

Buckner W., Reese E. & Reese R. (1997) Eye movement as an indicator of sensory components in thought. Journal of Counseling Psychology 34: 283–287.

Chalmers D. (2003) The content and epistemology of phenomenal belief. In: Smith Q. & Jokic A. (eds.) Consciousness: New philosophical perspectives. Oxford University Press, Oxford.

Costall (2006) "Introspectionism" and the mythical origins of scientific psychology. Consciousness and Cognition 15: 634–654.

Danziger K. (1994) Constructing the subject, historical origins of psychological research. Cambridge University Press, Cambridge.

Dennett D. (1992) Consciousness explained. Back Bay Books, Santa Ana CA.

Dennett D. (2002) How could I be wrong? How wrong could I be? Journal of Consciousness Studies 9: 13–16.

Depraz N., Varela F. J. & Vermersch P. (2003) On becoming aware. John Benjamins, Amsterdam.

Descartes R. (1990) Méditations métaphysiques. Livre de Poche, Paris. Originally published in 1641.

Ericsson K. A. & Simon H. A. (1984) Protocol analysis: verbal reports as data. MIT Press, Cambridge MA.

Garfield J. L. (1989) The myth of Jones and the mirror of nature: Reflections on introspection. Philosophy and Phenomenological Research 50: 1–26.

Horgan T. & Kriegel U. (2007) Phenomenal epistemology: What is consciousness that we may know it so well? Philosophical Issues 17: 123–144.

Hume D. (1978) A treatise of human nature. Oxford University Press, Oxford. Originally published in 1739.

Hurlburt R. T. & Heavey C. L. (2006) Exploring inner experience. John Benjamins, Amsterdam.

Hurlburt R. T. & Schwitzgebel E. (2007) Describing inner experience? Proponent meets skeptic. MIT Press, Cambridge MA.

Husserl E. (2002) Zur phänomenologischen Reduktion. Husserliana XXXIV. Kluwer, Dordrecht.

Husserl E. (2004) Ideas: General introduction to pure phenomenology. Routledge, London. Originally published in 1913.

Jack A. & Roepstorff A. (2002) The "measurement problem" for experience: Damaging flaw or intriguing puzzle? Trends in Cognitive Sciences 6: 372–374.

James W. (1890) The principles of psychology. Holt, Baltimore.

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

Kriegel U. (2013) A hesitant defense of introspection. Philosophical Studies. In press.

Lachaux J.-P. (2011) If no control, then what? Making sense of "neural noise" in human brain mapping experiments using first-person reports. Journal of Consciousness Studies 18: 162–166

Levi-Strauss C. (1958) Anthropologie structurale. Plon, Paris.

Locke J. (1975) An essay concerning human understanding. Oxford University Press, Oxford. Originally published in 1690.

Lyons W. (1986) The disappearance of introspection. MIT Press, Cambridge MA.

Mangan B. (2007) Cognition, fringe consciousness, and the legacy of William James. In: Velmans M. & Schneider S. (eds.) The Blackwell companion to consciousness. Blackwell, London.

Marcel A. (2003) Introspective report. Journal of Consciousness Studies 10: 167–186.

Marti S., Sackur J., Sigman M. & Dehaene S. (2010) Mapping introspection's blind spot: Reconstruction of dual-task phenomenology using quantified introspection. Cognition 115: 303–313

Merickle P. M., Smilek D. & Eastwood J. D. (2001) Perception without awareness: Perspectives from cognitive psychology. Cognition 79: 115–134

Nisbett R. E. & Wilson T. D. (1977) Telling more than we can know: Verbal reports on mental processes. Psychological Review 84: 231–259

Overgaard M., Koivisto M., Sørensen T. A., Vangkilde S. & Revonsuo A. (2006) The electrophysiology of introspection. Consciousness and Cognition 15: 662–672.

Petitmengin C. (2001) L'expérience intuitive. L'Harmattan, Paris.

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. & Bitbol M. (2009) The validity of first-person descriptions as authenticity and coherence. Journal of Consciousness Studies 16: 363–404.

Petitmengin C., Bitbol M., Nissou J.-M., Pachoud B., Curallucci H., Cermolacce M. & Vion-Dury J. (2009) Listening from within. Journal of Consciousness Studies 16: 252–284.

Petitmengin C. (2011) Describing the experience of describing? The blind spot of introspection, Journal of Consciousness Studies 18(1): 44–62.

Petitmengin C., Remillieux A., Cahour B., Thomas S. (2013) A gap in Nisbett and Wilson findings? A first-person access to our cognitive processes. Consciousness and Cognition 22(2): 654–669.

Piccinini G. (2003) Data from introspective reports. Journal of Consciousness Studies 10: 141–156.

Piccinini G. (2009) First-person data, publicity, and self-measurement. Philosophers Imprint 9(9): 1–16.

Pickering A. (1995) The mangle of practice. The University of Chicago Press, Chicago.

Prinz J. (2004) The fractionation of introspection. Journal of Consciousness Studies 11: 40–57.

Rudrauf D., Lutz A., Cosmelli D., Lachaux J.-P. & Le Van Quyen M. (2003) From autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being. Biological Research 36: 21–59.

Schwitzgebel E. (2004) Introspective training apprehensively defended: Reflections on Titchener's lab manual. Journal of Consciousness Studies 11: 58–76.

Shanon B. (1984) The case for introspection. Cognition and Brain Theory 7: 167–180.

Shanon B. (2003) The antipodes of the mind. Oxford University Press, Oxford.

Silverman M. & Mack A. (2006) Change blindness and priming: When it does and does not occur. Consciousness and Cognition 15: 409–422.

Thompson E. (2007) Mind in life. Harvard University Press, Cambridge.

Titchener E. B. (1912) The schema of introspection. American Journal of Psychology 23: 485–508.

Tye M. (2009) Consciousness revisited. MIT Press, Cambridge MA.

Varela F. J., Thompson E. & Rosch E. (1993) The embodied mind: Cognitive science and human experience. MIT Press, Cambridge MA.

Varela F. J. (1996) Neurophenomenology. Journal of Consciousness Studies 3: 330–349.

Vermersch P. (1994) L'entretien d'explicitation. ESF, Paris.

Vermersch P. (1999) Introspection as practice. Journal of Consciousness Studies 6: 15–42.

Wilson T. D. & Dunn E. W. (2004) Self-knowledge: Its limits, value, and potential for improvement. Annual Review of Psychology 55: 493–518.

Wittgenstein L. (1958) Philosophical Investigations. Basil Blackwell, Oxford. Originally published in 1953.

279