The overlooked ubiquity of first-person experience in the cognitive sciences

Joana Rigato[1] · Scott M. Rennie[1] · Zachary F. Mainen[1]

Abstract

Science aims to transform the subjectivity of individual observations and ideas into more objective and universal knowledge. Yet if there is any area in which first-person experience holds a particularly special and delicate role, it is the sciences of the mind. According to a widespread view, first-person methods were largely discarded from psychology after the fall of introspectionism a century ago and replaced by more objective behavioral measures, a step that some authors have begun to criticize. To examine whether these views are sufficiently informed by actual scientific practice, we conducted a review of methodological approaches in the cognitive science literature. We found that reports of subjective experience are in fact still widely used in a broad variety of different experimental paradigms, both in studies that focus on subjective experience, and in those that make no explicit reference to it. Across these studies, we documented a diverse collection of approaches that leveraged first-person reports, ranging from button presses to unstructured interviews, while continuing to maximise experimental reproducibility. Common to these studies were subjects acting as sensors, intentionally communicating their experience to the experimenter, which we termed "second-person" methods. We conclude that, despite views to the contrary, first-person experience has always been and is still central to investigations of the mind even if it is not recognized as such. We suggest that the conversation ought to be reframed: instead of debating whether to accept subjects' first-person knowledge we should discuss how best to do so.

Keywords First-person experience · Second-person methods · Subjective reports · Cognitive science · Epistemology · Phenomenology · Introspection

     Springer

1 Introduction: from first-person to third-person and back

In the early days of psychology in the late nineteenth century, introspection was valued as the direct access one can have to one's own mental states, which, in the words of William James, "we have to rely on first and foremost and always" (1890, ch. VII). In the initial development of experimental psychology, Wilhelm Wundt in Germany, followed by Edward Titchener in the United States, designed an extensive set of protocols in which the experimenter and subject exchanged roles and introspection was paired with careful measurements:

> A psychological experiment consists of an introspection or a series of introspections made under standard conditions. Some experiments are best performed by oneself on oneself. Most, however, require two persons for their performance: the observer O who makes the introspection, and the experimenter E who handles the instruments and makes the records. All experiments of this kind must be made twice over, O and E changing places. (Titchener 1901a, b, p. xiii)

However, early on, different labs endorsed incompatible epistemological assumptions (e.g. atomistic vs. holistic theories of the mind), which led to contrasting methodologies (e.g. standardized stimuli versus open-ended personalized questions) and conflicting results (e.g. an infamous controversy regarding the existence of imageless thought). While common textbook accounts of the history of introspection distort the facts by assuming mistakenly that the same experiments brought forth inconsistent results—an assumption that has plagued the reputation of introspection as a reliable method to this day—several authors have argued that the problems were actually grounded in conflicting theoretical views about the mind and the human ability to analyze it (Hatfield 2005; Feest 2012). Given these difficulties and the rising force of behaviorist research programs in the first half of the twentieth century, the introspectionist programme was considered largely defunct by the middle of the twentieth century.

The suspicion of "subjective" data was reinforced by a growing recognition in experimental psychology of the risks of experimenter biases, which may unconsciously influence both the subjects and the analysis. At the same time, it became clear that the behavior of experimental subjects is conditioned by their expectations about what is supposed to happen, as well as by the way they think the experimenter wants them to behave (Rosenthal and Rosnow 2009). In the process of designing protocols to avoid such risks, e.g. using groups of subjects rather than individuals or randomized and double-blinded trials, a clear separation between experimenter and subject was established as a crucial condition for the scientific credibility of psychological research. This is exemplified by this passage from Woodworth's influential Psychology textbook:

> [In] nearly all tests… the person tested is given a task to perform, and his performance is observed in one way or another by the examiner. The examiner may observe the *time* occupied by the subject to complete the task, or the *quantity* accomplished in a fixed time; or he may measure the *correctness* and

excellence of the work done, or the *difficulty* of the task assigned. One test uses one of these measures, and another uses another; but they are all objective measures, not depending at all on the introspection of the subject. (Woodworth 1921, p. 12)

These methodological improvements no doubt contributed to the overall reliability of psychological studies. Increasing emphasis was being placed on the principles of replicability and controllability in experiments which placed tight constraints on the types of behaviours, and experiences that one could expect to examine. Experiments were gradually defined as situations in which an experimenter controls the conditions under which an individual acts by manipulating one condition while holding all other conditions constant (Woodworth 1934). Thus, experimenters strived to constrain the range of behaviors analyzed in each task and trial, excluding from the list of analyzable phenomena all the subjective experiences that allegedly defied manipulation and quantification.

A turning point in psychology was reached by the middle of the twentieth century, when the limits of the behaviorist paradigm became undeniable and cognitive science, the analysis of the human mind in terms of information processing (Simon 1980), began its ascendance as the predominant paradigm in the sciences of the mind. Yet, while cognitive science brought mental phenomena and theories to center stage again, the experimental approaches retained their former character. These approaches continued to emphasize replicability, giving preference to controlling idiosyncrasies of individual subjects by methods such as randomized control groups, but also by constraining the dimensionality of experiential reports. This tended to limit first-person reports to non-verbal responses or to choosing from predefined categories of experience often in the form of Likert scales. By the 1980's the advent of powerful new techniques for noninvasive brain mapping including positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) launched a whole new era of cognitive neuroscience in which the same kind of psychological methods could be linked to putative neural substrates. Yet the approach to the characterization of mental processes themselves remained ostensibly third-person.

A few years later, philosopher Daniel Dennett introduced a method he termed heterophenomenology, which he announced as "*the* bridge—between the subjectivity of human consciousness and the natural sciences" (Dennett 2007, p. 249). On one side of this bridge would lie exclusively third-person data such as that regarding one's digestive tract or motor movements, i.e., facts that can be safely reduced to bottom-up causal explanations. On the other side would be private subjective experiences. The bridge would consist of the public reports of those private subjective experiences. While the experimenter could take these reports as a faithful transcript of *what it is like* for the subject to be in a certain phenomenological state, heterophenomenology cautions that she should remain agnostic about the *reality* behind the narratives.

However, it remains a matter of controversy just how far Dennett's proposal takes this skepticism. On a charitable reading, heterophenomenology only requires the experimenter to maintain the same healthy skepticism towards subjective reports

that one would wisely take to any scientific data. Just as electroencephalography must be considered an imperfect measure of neural activity, the reports of subjects must be taken as an "uncertain guide to what is going on in them" (Dennett 1991, p. 94).

Against this backdrop, there have always been self-proclaimed dissidents, those who insisted that the richness of subjects' internal lives was the most crucial and primary evidence to be taken into account in science. Some examples are found in Russell Hurlburt's decades-long development of the descriptive experience sampling method (Hurlburt and Heavey 2006a, b), K. Anders Ericsson and Herbert A. Simon's careful study of think-aloud protocols (Fox et al. 2011), or Pierre Vermersch's influential work on phenomenological interview techniques (1994). In the field of cognitive neuroscience, progress in the study of consciousness was grounded in the recognition of the crucial role subjective reports play and in the need to overcome "a century of behaviorist and cognitive suspicion" (Dehaene 2014, p. 46).

> In order to cross-correlate subjective reports of consciousness with neuronal or information-processing states, the first crucial step is to take seriously introspective phenomenological reports. Subjective reports are the key phenomena that a cognitive neuroscience of consciousness purport to study. As such, they constitute primary data that need to be measured and recorded along with other psychophysiological observations (Dehaene and Naccache 2001, p. 3)

Francisco Varela's neurophenomenology proposal (1996) aimed to shake things up by arguing for the need to revalue the subject's inner world and pair up Phenomenology with neuroscientific tools to investigate brain activity. Varela's proposal had a significant impact on the work of many cognitive scientists as well as epistemologists. Among philosophers, the question of the status of introspection as a source for evidence in cognitive science began to gain greater interest (e.g., Goldman 2004).

Since the 1990's, theoretical discussions about whether and how to "trust the subject" have been abundant. In 1999, the *Journal of Consciousness Studies* published a special double issue that resulted in the influential book *The View from Within: First-person Approaches*, edited by Varela and Jonathan Shear. In 2003, the same journal published another two-volume special issue, edited by Anthony Jack and Andreas Roepstorff, on "Trusting the Subject". In 2006 *Consciousness and Cognition* published a special issue on Introspection, edited by Morten Overgaard. In all three editions, numerous philosophers and cognitive scientists exchanged perspectives on the affordances and liabilities of using first-person experience as an evidential source of data in science.

We share with many of the authors that have discussed these issues in the aforementioned volumes the idea that science would benefit from overcoming its prejudice against subjective experience. However, we suspect that current debates about introspection may be out of touch with scientific practice. Rather than subjective experience being excluded as a source of scientific data, it maybe that it is simply not be being sufficiently acknowledged due to a lack of clarity in the manner in which it is defined. Psychophysics, for example, a branch of psychology that stemmed from the introspectionist tradition and that currently studies the relationship between physical stimuli and the mental responses they elicit, stands as a quite unproblematic

field where the perceptive experiences of single subjects are intersubjectively amplified to support generalizations about the human mind. Still, as Hatfield notes, "this general knowledge is to be achieved by introspective observations and their reports" (2005, p. 261). Could the use of first-person experience in science be more widespread that is usually assumed?

In order to verify what is the actual practice among scientists of the mind and brain, we decided to review the literature and create a taxonomic map of the various types of methods that contribute to the sciences of the mind: cognitive science, cognitive psychology, neuroscience, and so on. In order to do this, we gathered a comprehensive, although not exhaustive, sample of studies and rated the types of manipulations, neural recordings, behavioral measures and, of course, first-person reports that were used. We then tried to determine if there were any consistent patterns amongst these descriptors. As we suspected, our review confirmed that, despite cognitive science's ostensible third-person orientation, first-person experience is in fact widely used across the field in a variety of forms, despite not being recognized as such. Common to such studies was the *intentional use of communication between experimenters and subjects,* a feature we define as "second-person methods". We suggest this definition as it encompasses the diversity of ways in which first-person experience is captured. It emphasises that pressing a button in a psychophysics experiment still depends upon first-person experience, and should not simply be considered behavior, no less than verbal accounts.

2 A review of the cognitive science landscape

2.1 Disclaimer

When examining the literature regarding the use of subjective reports as evidence in science, one gets the impression of an unresolved tension between, on the one hand, claims that science should avoid using these reports altogether (following the influential 1977 article by Nisbett and Wilson) and, on the other, the growing recognition of the need to hone the best methods for a systematic use of introspective phenomenological reports in the study of conscious processes.

The goal of the quasi-empirical part of our research was to enrich our stance on this debate with a review of the methods that have been in use in the field. We do not claim to have presented readers with a complete map of the cognitive science landscape, nor to have analyzed the papers in our review with an exhaustive and fully replicable algorithm. Our only pretense is to have gathered an interesting collection of studies which have greatly contributed to illustrating the claims we will put forward in the third section of the present paper.

This paper is structured using the introduction/methods/results/discussion format, conventional in scientific literature, but less commonly found in philosophy journals. Hence, a reader wishing to understand immediately our argument and conclusions should just go to the Sect. 3. A reader interested in knowing the details of the papers in our sample and understanding the tools we used to analyze their features should go over the Sects. 2.2 and 2.3.

2.2 Methods

Our aim was to catalogue a diverse collection of studies in the cognitive sciences with respect to their experimental methodologies in order to see to what degree they relied on first-person experience and how this was related to other aspects of the methodology. We first present in detail the descriptors ('axes') we used to describe the experiments. We then explain how we selected the studies for inclusion. Finally, we explain how we clustered the studies into groups.

2.2.1 The axes

In order to systematically compare and contrast a diverse group of studies with heterogeneous and complex experimental methods, we constructed a set of 16 descriptors covering most aspects of the experimental methodology, from the way the subject reported their responses to questions to the type of neural recordings performed. Each descriptor was designed to be quantitative in nature, such that any experiment could be rated on a scale from 1 (low) to 10 (high) for each descriptor. Thus a descriptor corresponds to one axis in a 16 dimensional space and any study can be represented by a point in that space. This allows us to compare studies quantitatively, using the distance between them, and to apply other techniques to analyze the occupancy and structure of this space.

The specific axes we developed were as follows:

1. The first axis regards the *dimensionality of the reports* subjects use to intentionally communicate their first-person experience, which ranges from a binary yes/no button press (classified as 1) to unconstrained speech (classified as 10). Open answers to direct questions are rated with the intermediate classification of 5. In some cases, experiments included both simple button press reports, but also detailed unconstrained descriptions of the subjects' experiences during the task, in which case they would receive a high rather than low rating. The scale of this axis (as well as that of all the others) includes all the integers between 1 and 10, and when the scale is not applicable (which in this case means the protocols do not include reports of any kind), it is indicated by the classification of 0. This axis wishes to operationalize the question of the degree to which an experiment/method enables a subject to report on various qualities of her phenomenal experience.
2. Axis two, entitled *behavioral measures*, regards physical measures that are not under the explicit and conscious control of the subject (hence, a button press will not count as a behavioral measure), and not provided by recordings of the brain. They range from coarse behavioral measures such as posture (1), to more precise measures, such as an electromyogram (EMG) (3–5, depending on the number of electrodes), to fine grained psychophysiological measures (skin conductance, pupillary dilation, heart rate and so forth, rated 4–5), combined with motion tracking, which are rated from 6 to 10 according to their number and the information they provide. This axis addresses the question of how many different aspects of behavior each experiment/method collects data from.

3. With axis three, *neural recordings*, we rate the quantity and quality of the neural information gathered in each study. On the one hand, this depends on the average temporal and/or spatial resolution of the medium used, which can go from magnetoencephalography (MEG) (5), fMRI or PET scans (6), to single neuron recordings (9–10). On the other, it depends on the precision of the technological apparatus in use in the experience: an electroencephalogram (EEG), for example, can be can rated as 1 if it has only one channel, or as a 4 if it has 164. This axis operationalizes the question of how much information about the brain a certain experiment/method provides.

4. The fourth and fifth axes regard time. Axis four, *delay*, rates the time interval between the moment of experience and the moment of report (or measurement). For example, the two moments may be simultaneous (1), the report may be made immediately after the experience (2), or it may be made some hours or even years later, which implies the involvement of long-term memory (4–10). This axis operationalizes the question of how much retrospection is involved in a report.

5. Axis number five, *duration*, quantifies the duration of the experience that the method intends to capture. The possibilities range from the shortest instant that one can consciously perceive (1), to experiences that could even last several years (9). Cases when the subject generalizes from the sum of similar experiences to the consideration that she "always" feels a certain sort of pain or she is "usually" happy are rated as a 10. This axis assesses how elementary the experiences being reported are, while indirectly tackling the question of how much room for interpretation a method is likely to give to the subject.

6. In the sixth axis, we classify methods in terms of their degree of *reflectiveness*. Situations where the focus is mostly external to the subject, as for example attentional blink tasks where subjects are asked to identify a number among a sequence of letters, are classified as 1, more common cases where the subject is asked to report on the "quality of her inner experience" are rated with a 5, and situations where the object of the subject's attention is upon herself (e.g. personality surveys) are given a 10. With this axis we want to answer the question of how self-directed the introspective act is, in each experiment/method.

In most experiments, the subject is asked to achieve some goal, which may be more or less rigid and defined. This is done using certain means, which may also be more or less constrained. Axes seven and eight regard, respectively, the flexibility of the goal and the flexibility of the means. While the *goal flexibility* axis operationalizes the question of how limited the subject's options are among a set of pre-defined alternatives, the *flexibility of the means* axis regards the degree to which the protocol restrains the options of how to do it. In both these axes, we attribute the maximum grade (10) to the minimum level of constraint (i.e. highest flexibility).

7. According to the scale of the *goal flexibility* axis, a task that has an extremely constrained goal, as for instance a task where one has to reach a specific location in a computer game, is rated as a 1, whereas experiments in which there

is no predefined objective, such as simply exploring real world environments without a particular aim, would be a 10. The intermediate ratings correspond to tasks that are not fully defined, i.e. a task that allows for different alternative solutions (5).

8. In the *flexibility of the means* axis, we rate the degree of control the subject has over the means she may use in order to achieve the goal of the task. She may have no choice at all on how to achieve the goal, for example when simply being exposed to visual stimuli (1), she may have multiple options and strategies such as in competitive interpersonal games (5), or she may be allowed complete freedom of choice in how to reach a specific solution or goal (10).

9. In numerous studies reviewed here, subjects either underwent some procedure to manipulate neural activity or were explicitly selected from a clinical population where they suffered from known neural pathologies. The scale for *neural manipulation* axis refers to how specific these neural changes are. It ranges from very general manipulations such as those caused by the use of alcohol (1–2), relatively specific drugs or diseases (5–6), to very specific, fine-grained manipulations such as transcranial magnetic stimulation (TMS) (7) or the hypothetical use of optogenetics in humans (9–10). No manipulation is scored as 0. This axis operationalizes the question of how distant from a hypothetical baseline the brain state of the subject is.

10. Studies often included another type of manipulation that, instead of acting directly upon the brain, is better explained as a psychological process elicited by the experimental setting. In the *psychological manipulation* axis we address the question of the magnitude of these induced changes, by rating experiments according to the degree to which the protocol altered the mental state of the subjects, without any direct intervention at the neural level. Any task will be considered to induce a certain "executive mode" in the subject, which can be more or less close to her normal state (1–5). Deep changes due to the protocol, either self-induced (such as meditation) or induced by someone else (such as hypnosis), are rated according to their intensity (6–10). When protocols do not include any task or elicit any particular state via instructions, as in cases using ecological momentary assessment (EMA) techniques in real life, this scale is considered not applicable; hence they are rated as 0.

11. The various experimental approaches used in our sample vary not only according to the degree of constraints and manipulation, but also to what extent participants were informed of those constraints and the degree to which the experiment involved intentionally withholding information from them.

The scale of axis eleven, regarding the degree of *concealment*, ranges from being minimum (1) when the experimental manipulation, as well as the goals and constraints of the task are totally overt and the subject knows everything there is to know, to maximum (10) when the subject is completely deceived or completely uninformed. In the middle we rate the cases where a significant part of the variables is hidden but the experimenter does not give false information to the subject, or, vice versa, cases where there is a partial deception of the subject but most

variables are known (5). This axis operationalizes the question of the degree to which subjects are made aware of the experimental protocol.

12. Axis twelve has to do with the level of *ecological realism* of the experiment, i.e., how close the setup is to a natural environment. For example: sometimes experiments utilize a completely artificial context, such as the inside of a fMRI scanner (1), other times the experimenter tries to recreate a simulated environment, as for instance in classical psychophysics experiments (5), and others are conducted in an individual's every day environment (10). This allows us to indirectly tackle the question of how genuine one can expect the reported experiences to be, given the nature of the setup.

13. The thirteenth axis is *context complexity*. This measure concerns the complexity of the experimental setting, which ranges from the absence of any external input (1), to simple settings such as those of most perceptual tasks, in which the subject just has to look at an image (2), to increasingly complex settings which culminate in sophisticated virtual reality setups (9) or real life (10). This axis allows us to assess the spontaneity of the experience reported, and infer its complexity in stimulus-driven cases.

14. The fourteenth axis regards the degree to which the subject is in *control* of the stimuli she engages with. It ranges from total lack of control in setups where external stimuli, determined by the protocol, are presented to the subject (1), to meditation where the participant is engaging in a self-generated cognitive act (10). In between these two extremes we find various cases in which the subject has a variable degree of control over the setting, either because the stimuli come from the natural environment that she is dynamically engaged with (4), or because the experience is partially induced by an external stimulus and partially constructed by the subject herself, as in cases where a verbal input by the interviewer elicits a very complex cognitive process (6). With this axis we operationalize the question of the active role a subject has in the elicitation of the experience reported.

15. Axis fifteen concerns the amount of *training* that the subjects must have in order to perform the task. It ranges from no training at all (1), to some days (5), some weeks or months (6–7), or even several decades (10). This axis addresses the question of how much previous learning is required for the first-person experience in each method/experiment.

16. Finally, the *Subject N* axis reports the number of subjects involved in a typical instantiation of each protocol, ranging from one (1) to more than a hundred (10). The rating of N operationalizes the question of how idiosyncratic one can expect the reports to be.

2.2.2 The sample

We started collecting studies that we found to be most commonly cited among authors interested in neurophenomenology (e.g., Hurlburt 1997; Lutz et al. 2002; Ericsson 2003; Petitmengin 2006). We then expanded our focus to include studies that are currently published and discussed in conferences about these issues (e.g.,

Overgaard et al. 2006; Christoff et al. 2009; Ward et al. 2010; Garrison et al. 2013; Gallagher et al. 2015).

When we started reading and analyzing these studies, our set of 16 axes was not yet complete. It was early in the process of comparing them that the axes were fully defined. We then collected more studies concerning influential work in cognitive science that we judged to explicitly rely on first-person experience (e.g., Haggard et al. 2002; Libet et al. 1983; Penfield 1958, 1959; Carhart-Harris et al. 2016; Lahlou et al. 2015). Finally, we included in the review at least one representative exemplar of what we considered to be the distinct areas in cognitive science, from psychophysics, subliminal priming, work using virtual reality or video games, studies with brain damaged patients, studies using TMS or electric stimulation, to more esoteric examples of approaching conscious experience such as trip-reports (Erowid), and self-experiments. In total, we analyzed 53 studies or reviews. Some of these articles included multiple experiments with clearly distinct experimental methodologies, and in such cases they were analysed separately resulting in a total of 57 different methods (Fig. 1).[1]

To determine whether our sample could be considered representative of the breadth of methods used in the cognitive sciences we pseudo-randomly selected 20 further papers, which constituted our validation sample. We used an advanced Google Scholar search with the following criteria: that they have all three words 'subjects', 'experiment' and 'trial' in the text, and contain at least one of the words 'experience', 'interview', 'reports', 'questionnaire', 'cognitive', 'psychology', 'neuroscience'.[2] We used the first 20 publicly available papers, describing experimental methods and not only theory, returned by this search, and rated them (Fig. 2). We then calculated the minimum pairwise Spearman's rank distance of each study in the validation sample to any study in the primary test sample. The minimum distance of each method from the validation sample lay within the distribution of distances found within our test sample (Fig. 3). None of these new examples thus provided a novel or entirely unexplored collection of methods, suggesting that our sample was comprehensive.

It is important to note that the different studies we mapped were very heterogeneous in their representativeness. While some of them were quite unusual (e.g. Gallagher et al. 2015), others represented a whole tradition of research (e.g. Ratcliff and Rouder 1998). However, given the exploratory goal of this review, we considered this non-representativeness not to be problematic. We wanted to gather a panoramic perspective of a wide range of possibilities in the spectrum in order to develop a taxonomy of methods, not a faithful model of the field that could exhaustively account for their frequency.

---

[1] Four of these studies included different parts with distinct methods, and thus resulted in eight different items in our list.

[2] Query string: «subject, experiments, trial experience, OR interview, OR reports, OR questionnaire, OR cognitive, OR psychology, OR neuroscience».
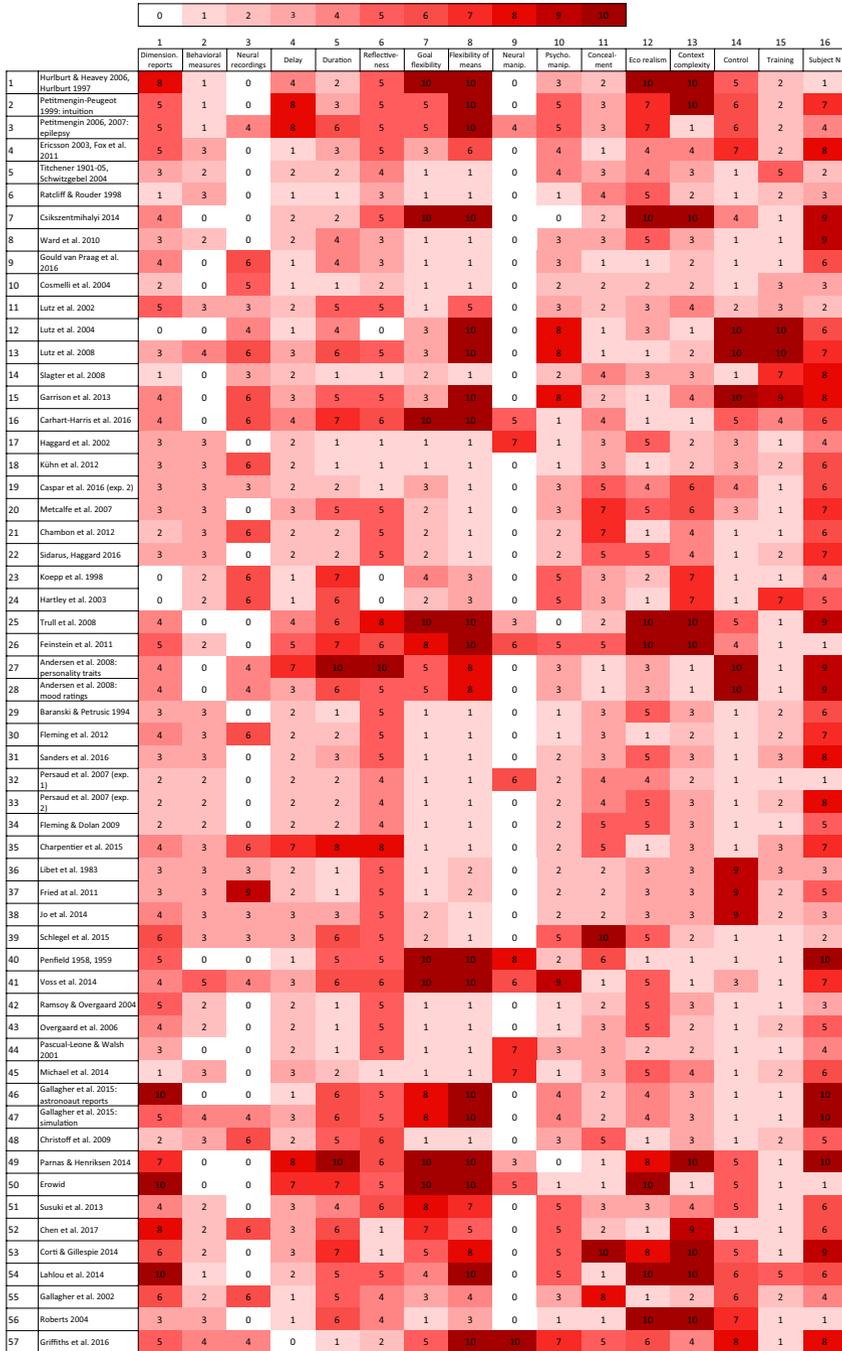
		1 Dimension. reports	2 Behavioral measures	3 Neural recordings	4 Delay	5 Duration	6 Reflective-ness	7 Goal flexibility	8 Flexibility of means	9 Neural manip.	10 Psycho. manip.	11 Conceal-ment	12 Eco realism	13 Context complexity	14 Control	15 Training	16 Subject N
1	Hurlburt & Heavey 2006, Hurlburt 1997	8	1	0	4	2	5	10	10	0	3	2	10	10	5	2	1
2	Petitmengin-Peugeot 1999: intuition	5	1	0	8	3	5	5	10	0	5	3	7	10	6	2	7
3	Petitmengin 2006, 2007: epilepsy	5	1	4	8	6	5	5	10	4	5	3	7	1	6	2	4
4	Ericsson 2003, Fox et al. 2011	5	3	0	1	3	5	3	6	0	4	1	4	4	7	2	8
5	Titchener 1901-05, Schwitzgebel 2004	3	2	0	2	2	4	1	1	0	4	3	4	3	1	5	2
6	Ratcliff & Rouder 1998	1	3	0	1	1	3	1	1	0	1	4	5	2	1	2	3
7	Csikszentmihalyi 2014	4	0	0	2	2	5	10	10	0	0	2	10	10	4	1	9
8	Ward et al. 2010	3	2	0	2	4	3	1	1	0	3	3	5	3	1	1	9
9	Gould van Praag et al. 2016	4	0	6	1	4	3	1	1	0	3	1	1	2	1	1	6
10	Cosmelli et al. 2004	2	0	5	1	1	2	1	1	0	2	2	2	2	1	3	3
11	Lutz et al. 2002	5	3	3	2	5	5	1	5	0	3	2	3	4	2	3	2
12	Lutz et al. 2004	0	0	4	1	4	0	3	10	0	8	1	3	1	10	10	6
13	Lutz et al. 2008	3	4	6	3	6	5	3	10	0	8	1	1	2	10	10	7
14	Slagter et al. 2008	1	0	3	2	1	1	2	1	0	2	4	3	3	1	7	8
15	Garrison et al. 2013	4	0	6	3	5	5	3	10	0	8	2	1	4	10	9	8
16	Carhart-Harris et al. 2016	4	0	6	4	7	6	10	10	5	1	4	1	1	5	4	6
17	Haggard et al. 2002	3	3	6	2	1	1	1	1	7	1	3	5	2	3	1	4
18	Kühn et al. 2012	3	3	6	2	1	1	1	1	0	1	3	1	2	3	2	6
19	Caspar et al. 2016 (exp. 2)	3	3	3	2	2	1	3	1	0	3	5	4	6	4	1	6
20	Metcalfe et al. 2007	3	3	0	3	5	5	2	1	0	3	7	5	6	3	1	7
21	Chambon et al. 2012	2	3	6	2	2	5	2	1	0	2	7	1	4	1	1	6
22	Sidarus, Haggard 2016	3	3	0	2	2	5	2	1	0	2	5	5	4	1	2	7
23	Koepp et al. 1998	0	2	6	1	7	0	4	3	0	5	3	2	7	1	1	4
24	Hartley et al. 2003	0	2	6	1	6	0	2	3	0	5	3	1	7	1	7	5
25	Trull et al. 2008	4	0	0	4	6	8	10	10	3	0	2	10	10	5	1	9
26	Feinstein et al. 2011	5	2	0	5	7	6	8	10	6	5	5	10	10	4	1	1
27	Andersen et al. 2008: personality traits	4	0	4	7	10	10	5	8	0	3	1	3	1	10	1	9
28	Andersen et al. 2008: mood ratings	4	0	4	3	6	5	5	8	0	3	1	3	1	10	1	9
29	Baranski & Petrusic 1994	3	3	0	2	1	5	1	1	0	1	3	5	3	1	2	6
30	Fleming et al. 2012	4	3	6	2	2	5	1	1	0	1	3	1	2	1	2	7
31	Sanders et al. 2016	3	3	0	2	3	5	1	1	0	1	3	5	3	1	3	8
32	Persaud et al. 2007 (exp. 1)	2	2	0	2	2	4	1	1	6	2	4	4	2	1	1	1
33	Persaud et al. 2007 (exp. 2)	2	2	0	2	2	4	1	1	0	2	4	5	3	1	2	8
34	Fleming & Dolan 2009	2	2	0	2	2	4	1	1	0	2	5	5	3	1	1	5
35	Charpentier et al. 2015	4	3	6	7	8	8	1	1	0	2	5	1	3	1	3	7
36	Libet et al. 1983	3	3	3	2	1	5	1	2	0	2	2	3	3	9	3	3
37	Fried at al. 2011	3	3	9	2	1	5	1	2	0	2	2	3	3	9	2	5
38	Jo et al. 2014	4	3	3	3	3	5	2	1	0	2	2	3	3	9	2	3
39	Schlegel et al. 2015	6	3	3	3	6	5	2	1	0	5	10	5	2	1	1	2
40	Penfield 1958, 1959	5	0	0	1	5	5	10	10	8	2	6	1	1	1	1	10
41	Voss et al. 2014	4	5	4	3	6	6	10	10	6	9	1	5	1	3	1	7
42	Ramsoy & Overgaard 2004	5	2	0	2	1	5	1	1	0	1	2	5	3	1	1	3
43	Overgaard et al. 2006	4	2	0	2	1	5	1	1	0	1	2	5	2	1	2	5
44	Pascual-Leone & Walsh 2001	3	0	0	2	1	5	1	1	7	3	3	2	2	1	1	4
45	Michael et al. 2014	1	3	0	3	2	1	1	1	7	1	3	5	4	1	2	6
46	Gallagher et al. 2015: astronaut reports	10	0	0	1	6	5	8	10	0	4	2	4	3	1	1	10
47	Gallagher et al. 2015: simulation	5	4	4	3	6	5	8	10	0	4	2	4	3	1	1	10
48	Christoff et al. 2009	2	3	6	2	5	6	1	1	0	3	5	1	3	1	2	5
49	Parnas & Henriksen 2014	7	0	0	8	10	6	10	10	3	0	1	8	1	5	1	10
50	Erowid	10	0	0	7	7	5	10	10	5	1	1	10	1	5	1	1
51	Susuki et al. 2013	4	2	0	3	4	6	8	7	0	5	3	3	4	5	1	6
52	Chen et al. 2017	8	2	6	3	6	1	7	5	0	5	2	1	9	1	1	6
53	Corti & Gillespie 2014	6	2	0	3	7	1	5	8	0	5	10	8	10	5	1	9
54	Lahlou et al. 2014	10	1	0	2	5	5	4	10	0	5	1	10	10	6	5	6
55	Gallagher et al. 2002	6	2	6	1	5	4	3	4	0	3	8	1	2	6	2	4
56	Roberts 2004	3	3	0	1	6	4	1	3	0	1	1	10	10	7	1	1
57	Griffiths et al. 2016	5	4	4	0	1	2	5	10	10	7	5	6	4	8	1	8

Fig. 1 Experimental sample composed of 57 studies (rows), rated along 16 axes (columns)

		Dimension. reports	Behavioral measures	Neural recordings	Delay	Duration	Reflective-ness	Goal flexibility	Flexibility of means	Neural manip.	Psycho. manip.	Concealment	Eco realism	Context complexity	Control	Training	Subject N
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
58	Kempf 1999 judgments about a product	4	0	0	3	7	3	5	4	0	1	2	6	6	3	1	8
59	Bridgeman et al. 1981 induced displacement	3	4	0	2	3	4	2	1	0	3	5	5	4	1	1	2
60	Greenwald et al. 1996 subliminal priming	1	3	0	2	1	1	1	1	0	2	7	5	3	1	1	8
61	Dehane et al. 1994 error-related negativity (ERN)	1	3	3	1	2	1	1	1	0	1	1	4	2	1	1	5
62	Marks 1973 visual imagery	4	3	0	3	5	1	2	2	0	2	5	5	4	1	2	8
63	Evans & Pezdek 1980 mapping real-world distances	1	3	0	1	4	1	1	1	0	1	4	5	3	1	2	6
64	Slovic & Fishhoff 1977 judgements of probability	5	0	0	1	0	1	2	10	0	1	7	6	10	1	1	10
65	Marcel 1983 unconscious perception	2	2	0	1	1	3	1	2	0	2	4	5	3	1	2	5
66	Lavie et al. 2004 attention and working memory	1	3	0	2	1	1	1	1	0	3	4	5	3	1	2	4
67	Hoyer & Brown 1990 brand awareness and choice	5	0	0	3	5	5	1	5	0	1	6	6	5	5	1	10
68	Notebaert et al. 2009 post-error slowing	1	3	0	2	3	1	1	1	0	1	5	5	2	1	1	5
69	Sternberg 1966 memory	1	3	0	2	3	1	1	1	0	2	4	5	3	1	2	4
70	Buckner et al. 1996 word-completion	0	0	6	0	6	0	1	5	0	3	4	1	4	1	1	2
71	Buxton & Meyers 1986 motor task w/ two hands	0	3	0	0	5	0	1	5	0	2	5	5	3	5	2	6
72	Smith & Swinyard 1983 attitudes vs. Behavior	4	2	0	2	6	4	2	2	0	3	6	6	6	3	1	9
73	Osu et al. 2004 motor task w/ force fields	0	3	0	0	7	0	1	2	0	2	5	5	4	1	2	4
74	Astafiev et al. 2004 EBA response to body movement	1	1	6	1	5	1	1	1	0	2	2	1	2	1	1	6
75	Fontani et al. 2005 effects of Omega-3	4	5	3	6	9	7	1	1	2	2	6	4	3	1	1	8
76	Phelps et al. 2001 threat, fear, amygdala	2	4	6	3	6	5	1	1	0	4	7	1	2	1	1	5
77	Eslinger & Damásio 1986 Alzeimer's and memory	3	3	0	3	5	1	1	3	6	2	3	6	4	2	1	5

Fig. 2 Validation sample composed of 20 pseudo-randomly chosen extra studies (rows), rated along the same 16 axes (columns)
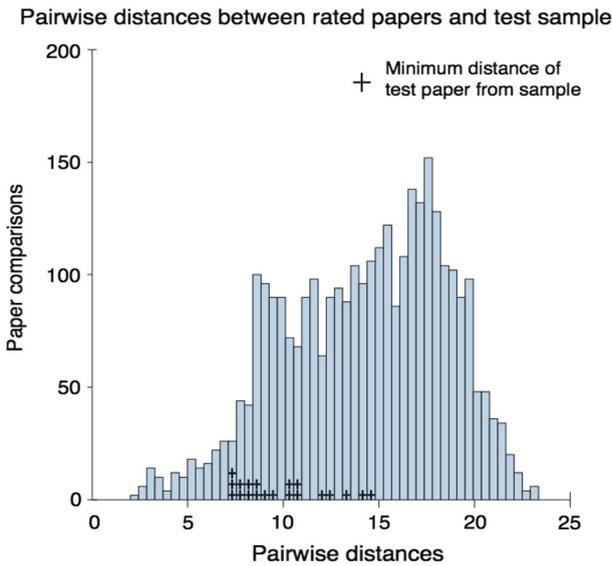
Fig. 3 Histogram of euclidean distances between each of the 57 rated studies in the experimental sample. Crosses show the minimum euclidean distances between each of the validation sample studies and the experimental sample

2.2.3 Clustering

We first calculated a Spearman's rank pairwise distance matrix for our rated sample of studies. This provided a measure of how different each sampled study was from

every other study based on their scoring across all 16 axes. We then performed a hierarchical clustering analysis with optimal leaf ordering and nearest neighbor linkage. This method clustered papers that were broadly similar to each other. All analysis was performed using MATLAB 2016a, The MathWorks, Natick, 2016.

2.3 Results

2.3.1 A matter of flexibility

It was evident from our analysis that a wide variety of methods are employed in cognitive science, most of which touch on first-person experience in some way—that is, they systematically involve a subject intending to communicate their experience to an experimenter. To begin to characterize the trends within this diversity of approaches, we first performed a hierarchical clustering analysis on the data (Fig. 4). This suggested that the studies fell into two broad clusters that correspond to, on the
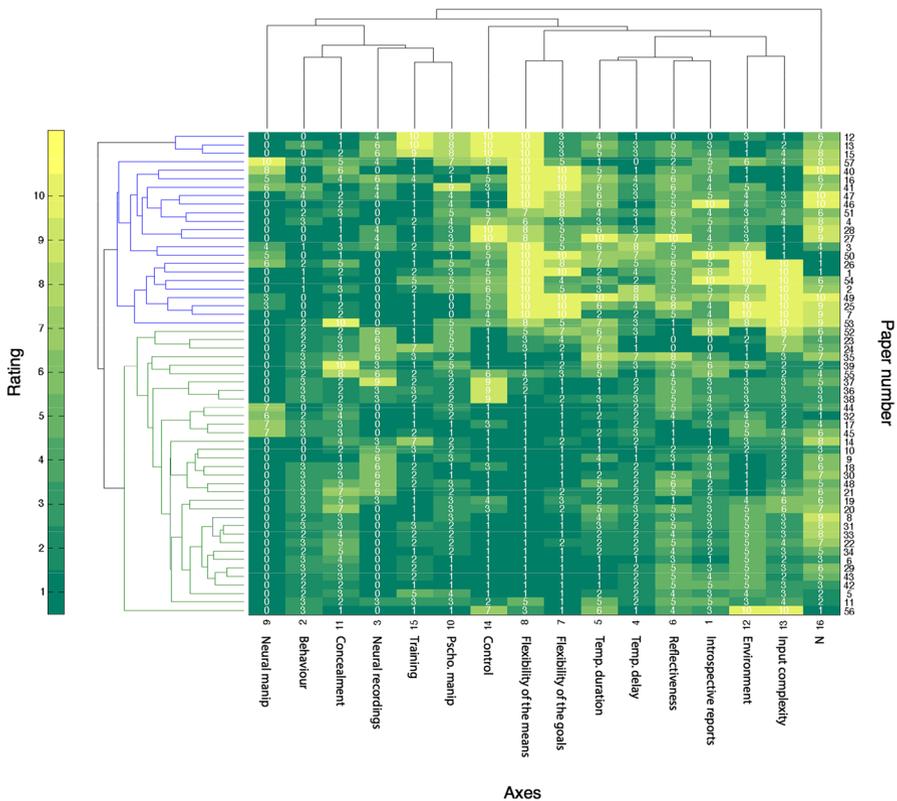
Fig. 4 A heatmap with dendrograms showing hierarchical clustering of each of the 57 papers (y axis) and each of the 16 rated categories (x axis) into clusters (branches representing flexible cluster in blue and constrained cluster in green)
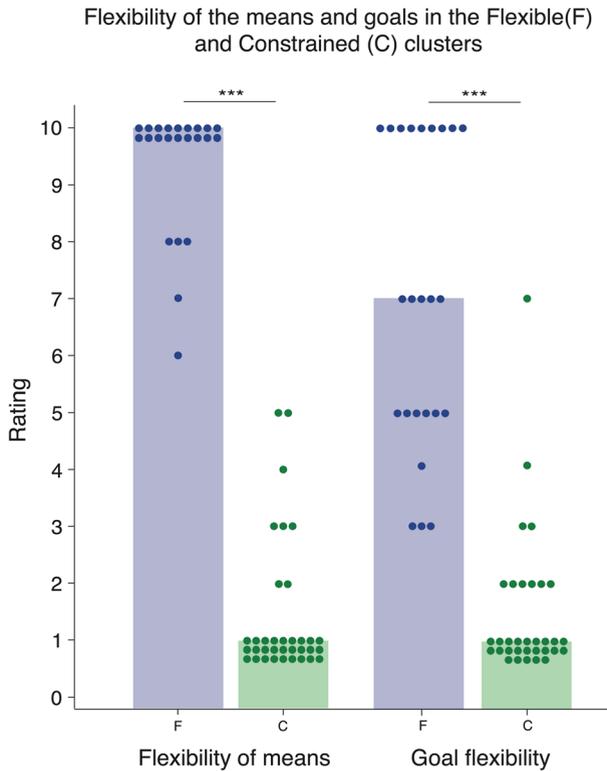
Fig. 5 Distributions of ratings (filled circles) and medians (bars) for flexibility of the means and goals for cluster 1 (blue) and 2 (green)

one hand, studies with high flexibility of means (axis 8) and high flexibility of goals (axis 9), and, on the other, studies with low flexibility of means and goals. This we confirmed by comparing the ratings for these axes between each putative cluster (Fig. 5). For this reason we have labelled the first cluster the *Flexible* cluster, and the second the *Constrained* cluster.[3]

The primary distinction between *Flexible* and *Constrained* studies can be described as how rigid and strictly defined (vs. flexible and loosely defined) the approach to collecting data from participants was. Supporting this distinction, studies falling within the *Flexible* cluster also tend to have higher ratings in control (axis 14) than studies in the *Constrained* cluster. This means that subjects in flexible paradigms tend to have more control over the input they encounter. In addition, studies in the *Flexible* cluster tended to also score higher in duration (axis 5) of reporting,

---

[3] The median values of both *flexibility of the means* (axis 8) and *flexibility of the goal* (axis 7) were significantly higher in the *Flexible* group of studies ($p < 0.001$, Wilcoxon signed rank test).
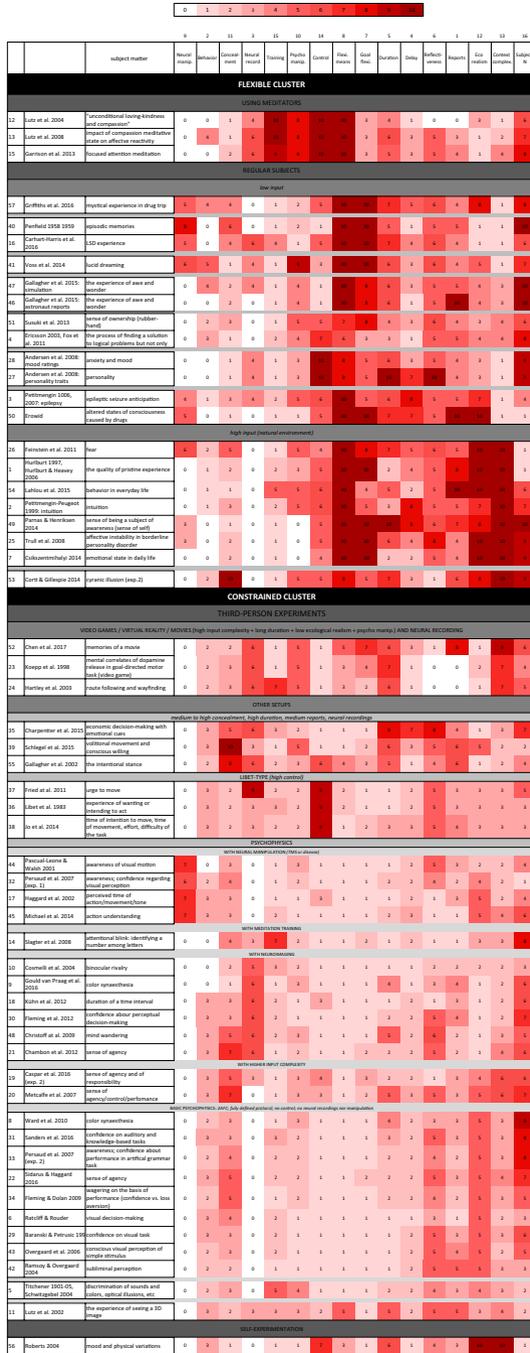
Fig. 6 Experimental sample re-organized according to the results of the clustering analysis

i.e. to report on longer experiences, e.g. several seconds-long experiences rather than short instants.

The hierarchical clustering analysis also suggested a number of potential subdivisions within the two main clusters, several of which we could easily identify and attribute meaning to (Fig. 6). In the *Flexible* cluster, one group of studies had in common the use of complex stimuli, usually in a naturalistic context, while a second group favored simple stimuli. One large sub-group of the *Constrained* cluster was particularly notable. It was defined by the traditional characteristics of psychophysical paradigms: two-alternative forced choice tasks, with a rigid protocol in which the subject has no control over a simple external input, and in which there are no neural recordings or neural manipulations. We also note variations on this typical format: studies using neural recordings; studies using very complex input (e.g. video games); studies using neural manipulation; studies where the subject has high control over the stimulus (e.g. Libet-type experiments, where the stimulus is the spontaneous "urge" to flex the wrist); and three studies where subjects do not report on their experience at all. In each of these three studies participants are asked to perform a task, compassion meditation in one (Lutz et al. 2004), and to play a computer game in the other two (Hartley et al. 2003; Koepp et al. 1998). In the two latter examples, aspects of the participants' gameplay provided performance metrics, and in the former, novice and experienced meditators were compared, allowing experience to be a proxy for meditation 'performance'. In each case these behavioural performance metrics were sufficient, alongside neural recordings, to test systems level hypotheses.

The methods used in both the *Flexible* and *Constrained* clusters include a variety of approaches that include neural recordings. One might expect that neural recordings would be more common in the *Constrained* cluster, however the frequency and spatiotemporal specificity in their use did not differ strongly between these clusters. The limitations of our sample mean that we cannot make conclusive claims about the prevalence of neural measurements alongside flexible protocols in cognitive science as a whole. Still, these data suggest that there is progress being made with neural data even with *Flexible* studies. The methodological challenges facing the analysis and interpretation of neural data within more flexible protocols appear to be in part addressed by combining expert subjects, in these cases meditators, who may be likely to have more consistent reports and behavior despite decreased experimental constraints placed upon them.

In general, while studies in both clusters collected a range of data regarding subjective experience, studies belonging to the *Flexible* cluster tended to use high-dimensional reports, while studies that fell into the *Constrained* cluster tended to use low-dimensional reports.[4] Interestingly, however, there were several studies

---

[4] A Pearson Rank correlation test showed a significant positive relationship between the flexibility of the means and dimensionality of reports ($r = 0.5$, $p < 0.001$, $n = 56$).

that used medium- to high-dimensional reports despite having been grouped under the *Constrained* cluster. Some of them included no explicit reference to the importance of leveraging the subject's first-person experience, e.g. Schlegel et al. (2015), a study using hypnosis to study conscious will in action, or Chen et al. (2017), an fMRI study about shared memories. Others were developed by experimenters that explicitly valued the use of introspective methods and were aiming at improving such methods, including Lutz et al. (2002)'s neurophenomenological study on the possibility of guiding the study of brain dynamics by using introspective reports, Ramsoy and Overgaard's (2004) study on subliminal perception, and Gallagher et al's. (2002) paper on the intentional stance. Importantly, when comparing the distributions of ratings for the dimensionality of phenomenological reports, we found that while they are higher on average in the *Flexible* than in the *Constrained* cluster, they clearly overlapped (Fig. 7).

Fig. 7 Distribution of ratings (filled circles) and medians (bars) for dimensionality of reports in flexibility cluster 1 (blue) and constraints cluster 2 (green)
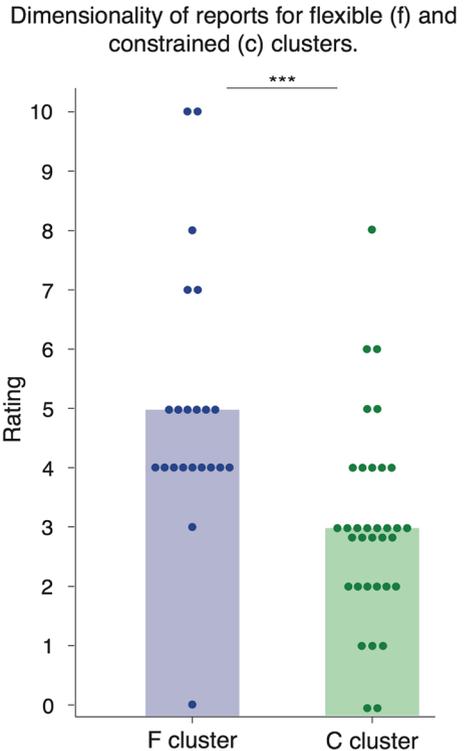
2.3.2 Experimental reproducibility

Increased flexibility—or decreased constraints—in performing a task provides opportunities for capturing a richer picture of subjective experience. Many studies in the *Flexible* cluster (and some in the *Constrained* cluster) used methods of reporting that allowed subjects to express themselves in free-form text or verbal responses. A central advantage in having such high-dimensional subjective reports was that there may be valuable nuance that could be gathered from intimate and detailed accounts of first-person experience. It is precisely this type of data that might otherwise be lost by experimenters limiting the dimensionality of reports *a priori* to simple Likert scales or behavioral responses.

At the same time, less constrained paradigms create challenges for experimental reproducibility, a concern that was at the heart of criticism brought against introspectionism. We can see two tactics taken to address this issue in contemporary studies. The first is to take steps to mitigate the increased variability that may be expected to ensue from decreased experimental constraints. Generally speaking there are two main sources of variability in typical cognitive studies: across-subject variability due to inter-individual differences, and within-subject variability due to uncontrolled variations across repeated tests. One way of dealing with inter-subject variability is to increase the number of subjects. Indeed, studies in the *Flexible* cluster tended to have a higher number of subjects than those in the *Constrained* cluster. However, increasing the number of subjects can only reduce variability if some form of averaging is applied—a step that may tend to obscure the very details that may be the aim of less constrained approach.

Another approach to variability is the use of expert participants. In Lutz et al. (2004, 2008), and Garrison et al. (2013), subjects were engaged in a meditation task with a specific goal (e.g. to achieve a state of unconditional loving-kindness and compassion) but were allowed high flexibility of possible cognitive means to complete this task. Here, the experimenters leveraged decades-long training in meditation of their experimental subjects, who were hypothesized to be more precise (less variable) at examining their own experience as a result of this training. This approach has been recommended by neurophenomenologists (Varela and Shear 1999a, b) as a way of reducing variability of first-person reports without constraining them a priori. While extensive training can be expected to reduce within-subject variability, training itself will not only make reporting more precise, but will often tend to mold experience itself in certain directions, thereby constraining it implicitly. Therefore, training in different schools of meditation might be expected to increase cross-subject variability even as it reduces within-subject variability in each individual or cross-subject variability within a school.

An alternative to training is to re-expose participants to the experimental tasks or conditions until the variability of their reports decreases. Arguably, this method has similar advantages to those of training whilst allowing for unanticipated aspects of experience to come to light. An important paper in our review that used this method is a study by Lutz et al. (2002) that is often cited as the quintessential example of the benefits of pairing phenomenological and neuroscientific tools (the "neurophenomenological paradigm"). Interestingly, this study fell within the *Constrained* cluster in

our analysis, the significance of which will be highlighted shortly. As this method is particularly well suited to highlighting idiosyncratic individual experiences, it does not address the problem of cross-subject variability and is better suited for within-subject comparisons.

While the approaches listed above aim to reduce uncontrolled variability in one way or another, it is also possible to take advantage of such variability rather than seek to reduce it. This is particularly applicable when seeking to establish a relationship between two measures, which would co-vary if they were indeed related to one another. For example, if the goal is to correlate subjective reports with neural measures, then temporal variability in a subject's state should impact both subjective report and neural measures. As described by Garrison et al. (2013):

> [We ask] participants to pay attention to how their own moment-to-moment experience changes, and to report how changes in their own experience relate to changes in the feedback graph. In this way, variability within individuals, and even variability within task blocks, does not confound results, but instead is utilized to more tightly couple subjective experience with brain activity (Garrison et al. 2013, p. 117).

A second tactic for dealing with problems of reproducibility in studies with high flexibility is the use of experimental interventions, such as drugs or electrical stimulation. Intervention studies are the hallmark of modern clinical trials and can make use of well-established paradigms to isolate the effect of a particular variable, the intervention, by comparing measures across "treatment" and "control" groups. Interventions can be combined with a wide range of task designs and degrees of constraint, and make a potentially powerful addition to studies using flexible reporting. For example, Voss et al. (2014), used a within-subjects comparison to study the impact of electrical stimulation on lucid dreaming. During REM sleep, subjects underwent fronto-temporal transcranial alternating current stimulation (tACS) at various frequencies, alongside simulated stimulation, with no current flow. After the stimulation, they were awakened and asked to report on the quality of their dreams. The within-subject comparison between different conditions (different frequencies as well as sham stimulation) allowed for the isolation of the causes of the variations detected by the subjects in their own dreaming experience. Carhart-Harris et al. (2016), studied the phenomenology associated with LSD using a blinded, placebo-controlled, across-subjects comparison: both treatment and control groups were prepared as if they were receiving the drug, but only one group actually received it. Thus, while "trusting the subject" to describe their experience, this design helped to prevent subjects from unwittingly using their knowledge of the study to bias their descriptions. The experimental design also used within-subjects comparisons between the reports collected at three different time points post LSD and placebo injection as a way to better isolate the perceived effects of the substance.

Blinding is a particularly powerful experimental method that is also applicable to many other facets of experimental design. For example, in Gallagher et al.'s (2015) study of astronaut experiences, hermeneutical analysis of diaries was performed by those blind to the examined hypotheses to reduce the possible effect of experimenters being biased in their interpretation of the data.

2.3.3 Triangulation

In discussing the role of introspective methods in contemporary cognitive science, several authors have argued for the use of triangulation as a potentially powerful approach in advancing neurophenomenology and the understanding of first-person experience in general (Jack and Roepstorff 2002; Gallagher 2002). Triangulation is typically defined as the simultaneous collection of (1) subjective reports, (2) neural recordings, and (3) behavioral measurements. Surprisingly, even though a variety of papers amongst those we assessed here were strong in two of the three components, none were highly rated in all three.

This suggests that, while triangulation may be a powerful approach, it may be challenging to implement or it may still not have gained sufficient traction in the field. Although behavioral measures should on the whole be simpler to implement than neural recordings, we noted a relative dearth of highly quantitative behavioral methods, such as eye tracking, in the sampled studies. When they were present, they were very low dimensional (usually limited to levels of performance or reaction time), especially in the *Constrained* cluster where they were never rated higher than 3 out of 10 on our scale. In the *Flexible* cluster, behavioral measures were also often absent, but in four out of the 24 studies (17%) we gave ratings of 3 or 5.

Despite the scarcity of behavioral data, we found several studies that at least come close to triangulation: Charpentier et al. (2015), Gallagher et al. (2002, 2015), Fleming et al. (2012), Jo et al. (2014), Lutz et al. (2002), Schlegel et al. (2015) and Voss et al. (2014). Interestingly, they were equally likely to be in the *Constrained* or *Flexible* cluster.

2.3.4 Different types of experiences

Finally, our review shows a clear and important difference between the two clusters in the types of experience they take into account. In the *Flexible* cluster, papers are mostly dedicated to experiences that are currently at the edge of explanation and that have not yet been, and might not easily be, reduced to more operationalized behavioral responses. For example, experiences such as those elicited by psychedelic trips, or the awe of seeing the Earth from space, might not be straightforwardly reflected in the behavior of subjects nor adequately captured by low-dimensional reports using predetermined Likert scales or button presses. Conversely, the *Constrained* cluster contained studies focused mainly on experiences that tend to be expressible behaviorally, or have been previously operationalized, such as perceptual experience, decision confidence or loss aversion, which have been carefully shown to be highly correlated with third-person measures such as, respectively, performance in a detection task, post-decision waiting time, or a characteristic wagering pattern.

3 Discussion: the presence of first-person experience

3.1 Summary

The place of subjective experience in scientific explanation is commonly viewed with suspicion, even in sciences purporting to deal with the most intimate and subjective aspects of individual experience. Accordingly, it is seldom discussed or acknowledged in the mainstream cognitive sciences. In recent years, an opposition to this view has arisen, with a number of authors strongly insisting on the value of first-person experience and arguing forcefully against it being largely excluded from cognitive science. We agree that first-person experience is fundamental to cognitive science and allied fields such as cognitive neuroscience. However, we contend that, far from it being excluded from these sciences, it is in fact nearly ubiquitous within them, yet largely unacknowledged. We support this view using a review of the landscape of current research. We quantified a sample of 53 selected studies from cognitive science literature. This sample was chosen to include as broad a range of experimental approaches as possible amongst studies purporting to address some aspect of human cognition. Consistent with our hypothesis, we found that the large majority of studies required subjects to report on their first-person experience. What varied across these studies was the degree of freedom in the report, from psychophysics (very limited), to semi-structured interviews, to self-report questionnaires (freeform). A small minority of studies did not include subjective report—specifically Lutz et al. (2004), Koepp et al. (1998) and Hartley et al. (2003). Thus, first-person approaches were not required in the cognitive sciences, but were far from being neglected.

To try to understand better what distinguished studies considered (by philosophers) to be "first-person" from others, we classified each of 57 different experiments in these studies along 16 quantitative descriptors (see Sect. 2.2.1) corresponding to different characteristics of the experimental approach. We aimed to capture the nature of the information gathered in each experiment (which type of subjective reports, behavioral measures and neural recordings were used), the kinds of experience targeted by these measures, and the characteristics of the task performed by the subjects, as well as of the experimental setup.

This revealed that the studies surveyed clustered into two main groups: one we labelled the *Flexible* cluster, which comprised studies that typically allowed subjects substantial freedom, and the *Constrained* cluster, which grouped studies where subjects had a more limited set of behavioral alternatives. *Flexible* cluster studies tended to collect higher dimensional reports than those in the *Constrained* cluster, but the distinction between the two main clusters was not primarily grounded in the reporting methods, rather in the constraints of all kinds imposed on behavior by the task—such as limited head movement, two-alternative forced choice tasks, etc. Moreover, only authors of studies in the *Flexible* cluster typically acknowledged the use of first-person data.

Thus, we conclude that first-person experience is being explored with a variety of different methods in the cognitive sciences, and that the debate in philosophy

regarding the place of phenomenology in cognitive science is to some degree out of touch with experimental reality. We suggest that once the ubiquity of first-person experience in cognitive science is acknowledged, a main challenge for the field, together with philosophy, will be to thoughtfully utilize the space of alternative methodological approaches in order to help bring the richness of subjective data into scientific light. We hope that our analysis has contributed to revealing how large the space for methodological exploration is.

3.2 First-person experience is conveyed via second-person methods

"First-person experience" is conventionally defined as the subjective and qualitative phenomena that constitute the inner world of an individual, the what-it-is-likeness to be that individual. In contrast, "third-person observations" conventionally concern behavioral or physiological phenomena that are externally measurable by observers and are hence "objective". Third-person data may include volitional responses such as button presses or facial expressions as well as processes that subjects do not even voluntarily control, such as skin conductance, neural activity or reaction times. Even with a commitment to brain-mind identity, the explanatory gap between these subjective and objective perspectives can seem dauntingly wide. How could an external observer possibly link an objective behavioral measure to any first-person experience, to which, by definition, she has no access?

Yet, as we examine the long list of experiments in our review, we find in all of them a common approach to bridging the gap between first-person and third-person, which is the *intentional use of communication between experimenters and subjects.* It is precisely through such acts of communication that a willing subject can convey her inner experience to another person. We propose that the methods in which acts of communication are used to provide data about first-person experience be named "second-person methods".

Figure 8 illustrates this terminology with an example from psychophysics. The example was adapted from Cosmelli et al. (2004) and depicts a subject in a binocular rivalry task. Here, we can identify three levels of description: the level of first-person experience, the level of second-person report, and the level of third-person data. First, there is the first-person experience itself of seeing a given stimulus at a given time t (1). The subject expresses this experience in the form of a second-person report, which in Cosmelli et al.'s experiment is pressing one of two buttons to indicate the presence or absence of the stimulus, but could also have been a verbal 'yes' or 'no' response (2). Finally, the experience can be indirectly accessed from a third-person perspective through physiological and behavioral data, such as neural recordings or pupillary dilation (3).

The main advantage of assuming a second-person terminology is that it allows us to distinguish these three levels very clearly. If we consider subjective reports as first-person data, we are not taking into account the difference between what one feels and what one says she feels, which is a mistake. Only the subjects themselves have direct access to their subjective experience, despite their intention and

Fig. 8 Illustration of an experiment (adapted from Cosmelli et al. 2004) where the first-person experience is conveyed via second-person methods and captured indirectly through third-person measurements
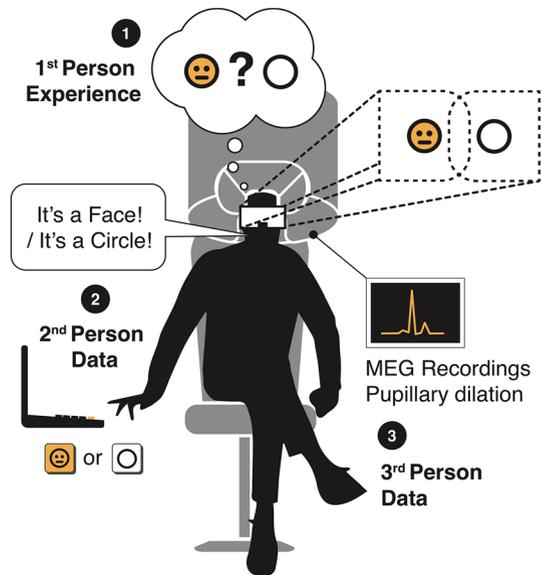
sometimes trained and highly skilled competence in translating it into a report. By second-person reports, we mean the utterances, reports, button presses, and other public objects that constitute the data on which cognitive science can be based (Piccinini 2009). On the other hand, though, if we consider such reports as third-person objects, we do not take into account the difference between data that depends on what the subject feels and translates into words—hence data that depends on her ongoing experience—and data that does not. The second-person terminology allows us to capture both these differences.

Crucially, what distinguishes second-person reports from mere third-person observations is the *intention* of the subject to convey her experience, together with the "intentional stance" (Dennett 1996) whereby the experimentar interprets her utterances or gestures as words and symbols. An objective button press by a subject is linked to that subject's subjective experience by her understanding of the meaning that the experimenter assigns to the action and her intention to translate her phenomenological experience into this category. The same button press by a subject who does not wish to articulate her inner experience or experiences a lapse in the ability to do so remains an objective measure but is no longer a second-person report.[5]

---

[5] Note that the sense in which we here use the phrase "second-person methods" is distinct from the stricter use Varela and Shear (and their followers) give it in (1999b, b). They refer to interview methods in which an expert mediator is required for the subject to be able to notice her own experience and put it into words. We instead use the term "second-person" to label all the reporting methods whereby a subject communicates aspects of her subjective experience to another person. It is also different from the sense in which Andreas Roepstorff speaks of the "second-person perspective", which pertains to "the fact that so much of human consciousness and perception is directed against and mediated by inputs from other people" (2001, p. 762). However, our account of second-person methods is very close to Jack and Roepstorff's call for a reconsideration of the "communicative interaction" that is present "both when the experimenter gives the instructions and when the subject describes their experience" (2002, box 1).

Second-person reports are not always accurate. Errors can be reduced by having trained or otherwise skilled participants, but subjects may inadvertently fail in their attempt to identify and/or understandably communicate their mental states. In that case, the second-person data provided by their reports will be inaccurate. This is something that the subject may or may not be able to indicate, depending on whether she is aware that her description is incomplete. But as long as subjects are not intentionally lying, the epistemological status of their reports remains unchanged, as they are the result of the intentional communication of information by the subject regarding her first-person experience.

The distinction between second- and third-person data is clear enough in the case of verbal reports, because verbal (or written) communication is nearly always conscious and intentional. In the case of nonverbal reporting, the distinction can be more subtle. An experimenter cannot distinguish with absolute certainty whether a subject's button press signified a meaningful attempt to interpret inner experience or instead resulted from deliberate lie or careless movement. Experimenters must therefore rely in part on other aspects of the experimental methodology to help interpret the epistemological status of such data. The subject herself often has a opinion about whether she was in fact conveying some meaningful information about her first-person experience, so-called metacognitive access, which itself can be queried using scales or button presses. Experimenters may also make use of other cues such as the consistency of reaction times or the correlation between nonverbal reports and other behavioral or physiological measures. A verbal debriefing after the experiment is often used to ascertain whether the subject's experience and reporting was in accord with what the experimenter expected.

3.3 Second-person reports are not necessarily introspective

One of the merits of the second-person terminology is that it makes the continuity between low- and high-dimensional reports clearer. From the moment we understand that methods apparently as distant as psychophysics and Descriptive Experience Sampling are actually equally bound by the intentional communication between subject and experimenter, thus belonging in the same category, this terminology becomes the most adequate. It includes studies with both high and low dimensionality of reports, that use verbal as well as non verbal tools as a way to report on the subject's experience.

However, one may wonder whether the "second-person" terminology is a way to shy away from using the term "introspection" to avoid the risk of losing grip on this notoriously difficult terrain. While we assert that psychophysics is grounded in the use of second-person methods, there remains some question as to whether low-dimensional reports such as button presses are the result of an "introspective" process or not.

There are many definitions of introspection. In a broad sense, introspection is understood as "deliberate and immediate attention to certain aspects of phenomenal experience" (Hatfield 2005, p. 279). According to this perspective, any second-person report is an introspective report. In a more restrictive sense, however, introspection can be defined as the act of reflecting upon one's phenomenal experiences—"a

reflective second-order cognitive act that thematizes first-order phenomenal experience, and makes that experience the object of reflection" (Gallagher and Overgaard 2006, p. 278). According to this definition, introspection occurs only when the subject *reflects* upon her experience, which is different from immediately attending to it. To use the common "inner observation" metaphor, introspection as attention is similar to the casual look with which one sees the world as one goes about doing other things, whereas introspection as reflection is rather like the systematic observation required when one wishes to seize an image in detail. Personality tests and self-assessment methods—whereby subjects are asked to examine their own states of mind, emotions, or mood—are typical methods involving reflective introspection. But many experiments in cognitive science do not obviously require the use of reflective introspection. In the binocular rivalry experiment, for example, the image seen is a subjective experience on which individuals are asked to report via button presses. When they press a button indicating that they are seeing image A rather than B, they are reporting on a concurrent first-order phenomenal experience (seeing the image). According to Gallagher and Overgaard's perspective, this proceeds without the need for a second-order introspective reflection about that experience:

> We can report on what we experience without using introspection because we have an implicit, non-introspective, prereflective self-awareness of our own experience. [In an experiment where I'm asked to press a button when I see a light come on,] at the same time that I see the light, I know that I see the light. This knowledge of seeing the light is not based on reflectively or introspectively turning our attention to our own experience. It is rather built into our experience as an essential part of it, and it is precisely that which defines our experience as conscious experience. (Gallagher and Overgaard 2006, p. 279)

Conversely, one can also easily imagine scenarios in which the result of a complex and painstaking introspective process could be reported using a button press. Indeed, it may be possible to access the very same kind of subjective experience using introspection or not, depending on task instructions and perhaps training, and have it reported using the same sort of button press. For example, one thing is to report on what one sees (perception) and another is to report on the experience of seeing (introspection). Using the broader definition of introspection as deliberate attention, a given task could be performed with lesser or greater introspective skill. Slagter et al. (2008), for example, documented the impact of a 3 month Vipassana meditation training in the performance of subjects in an attentional blink task. After this training, subjects' showed an improved capacity of "being attentive moment by moment to anything that occurs in experience, whether it be a sensation, thought or feeling". On the basis of these results, we would expect that had these subjects been tested before and after on a given perceptual task it could have been performed more introspectively after the training.

According to Mazviita Chirimuuta, we may avoid the difficult task of finding a consensual definition for introspection but still distinguish between "minimally-introspective" and "introspection-heavy" tasks in the well-known terrain of psychophysical experiments. The former are tasks in which subjects are asked to report on the presence/absence of the stimulus or on its apparent properties (e.g. color, movement),

while in the more introspection-reliant tasks, subjects engage in less immediate judgments, such as "careful comparison of sensory experiences that bear non-obvious relationships of similarity and difference to each other" (Chirimuuta 2014, p. 918). This is not an easily quantifiable distinction, but such an intuitive qualitative difference is corroborated by similar dichotomies found within the psychophysics tradition, which sometimes distinguishes between "class A" and "class B" experiments. In "class A", the experimental subject is treated like a "thoughtless measuring instrument" while in "class B" she is expected to be "a critical being who can attend to and reflect on her own conscious states" (Chirimuuta 2014, p. 922).

Our survey includes several methodological descriptors that may be useful indicators of the type or degree of introspection used in a given case. Our fourth descriptor, "delay", measures the time interval between experience and report, which is likely to be correlated with the degree to which a subjective report relies on a time-consuming process of introspection. A very short time interval between experience and report will prevent subjects from introspecting deeply, while when there is more time available for reflection, introspection is possible and perhaps even encouraged. When the distance between experience and report is too large, however, introspection will arguably give way to retrospection or recollection. The sixth descriptor, regarding the degree of "inward reflectiveness" of the protocol, can also provide indirect information about their introspective-ladenness. If the reports regard explicitly the inner experience of the subject, such as feelings or thought processes, they will most likely rely on her introspective skills. On the other hand, if the object of attention is external to the subject, even though it still regards inner experience, introspection may be expected to be less engaged in the process leading from experience to report, unless explicitly directed.[6]

---

[6] We are aware there are more radical positions where introspection may be much less involved in metacognition than is commonly assumed. One such position is Peter Carruthers' theory about the lack of introspective access to judgments, decisions and the like, according to which "our knowledge of our own attitudes results from turning our mindreading capacities upon ourselves" (Carruthers 2009, p. 121). In other words, our knowledge of our beliefs and intentions stems from an unconscious interpretative process based on perceptual or quasi-perceptual cues, rather that on a putative direct access to phenomenal experience. We believe that an assessment of this kind of theory is orthogonal to the scope of the present paper, although we are sympathetic to rival positions, such as Alvin Goldman's, according to which, even if people sometimes confabulate when the introspective access to their propositional attitudes, for some reason, is defective (e.g. split-brain subjects), the default access method we use to become aware of our mental states in normal situations is still introspection, which should be accepted as a good evidential source, to be used under operational conditions that can be easily identified, allowing us to agree upon a certain "range of introspective reliability" (2004, p. 14). The question some may ask is: what would the consequences be if one were to follow Carruthers' theory instead? Under his higher-order theory of consciousness, according to which "it is the availability of globally broadcast states to the mindreading faculty that is responsible for their phenomenally conscious status" (p. 124), introspection is a necessary condition for consciousness, which means that the absence of the former would entail the absence of the latter. We are not committed to such theory, though. According to our perspective, consciousness is a precondition for introspection, but the reverse is not true. Even if our cognitive access to a subset of our mental states, such as our propositional attitudes, is not direct nor transparent, there is still something it is like to experience such beliefs, intentions and judgments. The metacognitive access subjects have to their mental states may be the result of some sort of interpretation, but they are still experienced as part of the subject's phenomenal world. Also, even if part of our mental states are unconscious, that does not invalidate the fact that subjects are able to report on the conscious states they effectively have.

It is evident that introspection needs to be better understood both phenomenologically and neurally and that its precise definition remains elusive. In contrast, what constitutes second-person data is clear. Whenever a subject becomes part of the experimental setup as a detection or measuring device (Piccinini 2009) of her first-person experience, the data becomes second-person. The validity of second-person data collected depends not on the degree of introspection but on the ability of subjects to understand what is asked of them and to intentionally convey in a meaningful way their subjective experience to another person, even when that experience is immediate and their self-awareness is prereflective. And so while there may be ongoing terminological discussion as to what introspection is, all sides of the discussion will likely agree that all data gathered from introspection can be considered what we term second-person. Whether introspection was employed, and in which manner, in the collection of second-person remains an important question, but a secondary one.

3.4 First-person experience is everywhere

It has been prominently argued in the literature that first-person approaches (what we have termed second-person methods) are required for the study of subjective experiences that seem not readily detectable using behavioral measures. These include experiences such as dreams (Voss et al. 2014), visual imagery (Marks 1973a, b), synesthesia (Ramachandran and Hubbard 2001), psychedelic trips (Carhart-Harris et al. 2016), spiritual experiences (Lutz et al. 2008) and ineffable emotions (Gallagher et al. 2015). For these sorts of cases, it is easy to assume that research based on less-constrained linguistic reporting is likely to be needed before an element of the experience can be distilled to a binary or scalar quantity. Personality tests and self-reported psychiatric diagnostic instruments, such as the Beck Depression Inventory (Beck et al. 1961), are prominent examples in which questions of an extremely subjective and introspective nature have been translated into multiple-choice answers. While one may debate the validity of such measures, it is apparent that the dimensionality of reporting may be reduced arbitrarily far, so long as sufficient care has gone into ascertaining their meaningfulness for subjects.

However, as we have argued above, it is not only complex introspection-laden topics such as these that require the use of first-person experience. Even in experiments based solely on third-person data, i.e. when the primary behavioral or physiological data is available to collect even regardless of the subject's intention to report it, subjective aspects of experience may still play an important role. The fact that an unambiguous and imminently quantifiable behavior such as a button press can nonetheless be considered an intentional report of a subjective state, similar to a verbal response in an interview, has seldom been recognized in the literature. Even the adversaries of an overly reductionist cognitive science often assume that button presses are "objective behavioral responses" (Cf. Jack and Roepstorff 2002, Box 1), as opposed to subjective reports. However, when we analyze the various means available to a subject in order for her to report her first-person experience,

from yes/no button presses, to numerical scales and multiple choice questions, to open-answer questions, to semi-structured interviews or diary entries, we find that all of these reporting methods can be used by subjects to voluntarily provide the experimenter with information about their subjective experience. The methods vary in the number of dimensions they include, which increase exponentially as subjects are given freedom to express themselves with their own words rather than via pre-defined categories of experience. Still, these differences regard only the complexity and tractability of the data, not their epistemological status. It is an illusion to think that scientific experiments on mental processes can banish the issues regarding the evidential status of these reports simply by reducing the dimensionality of the response to a binary choice.

Besides the reports themselves, first-person experience and second-person methods (intentional communication) are essential through many other aspects of the broader experimental context. As Dennett himself acknowledges, this "practice of talking to subjects (…) is an ineliminable element in psychological experiments" (Dennett 1991, p. 74) To design tasks, scientists must understand to some degree the experience their subjects go through, which in part relies on using their own experience in similar contexts. On the other hand, subjects are typically expected to understand through verbal instructions what experimenters want them to do: the goals and rules of the task, and their understanding depends on a shared language through which subjective experiences are translated into words and can thereby be discussed publicly. According to Max Velmans, this exchange allows a transition from subjectivity to intersubjectivity to occur, since, "through the sharing of a similar experience, subjective views and descriptions of that experience potentially converge, enabling intersubjective agreement about what has been experienced" (Velmans 1999, p. 304).

Although humans can obviously perform tasks with no verbal instructions, just as animals must do, this is rarely done. And even when purely behavioral measures are used as proxies for cognitive states as, for example, pupillary dilation for surprise (Preuschoff et al. 2011) or post-decision wagering for confidence (Persaud et al. 2007a, b), they must have been previously validated through second-person reports that correlate them with first-person experience. In our sample, only three studies did not directly leverage subjective experience or second-person methods in their experimental design. For example Lutz et al. (2004) examined the relationship between compassion meditation in long term meditators and gamma synchrony. The experimental design did not require the participants to communicate with the experimenter, but simply to engage in meditation. Similarly Koepp et al. (1998) examined striatal dopamine release in participants as they engaged in a computer game; participants simply played the game and again were not required to report on aspects of the gameplay to the experimenter. However, the choice to examine dopamine release during computer game play, or gamma activity during meditation did not come about in the absence of knowledge gained from first-person experience and second-person methodologies. For example, the choice to examine gamma band oscillations via EEG was informed by a body of knowledge previously built from descriptions of the mental strategies engaged by meditation practitioners and their subsequent experiences.

Similarly, the choice to use a certain type of video game has likely relied on hypotheses influenced by the personal experience of the experimenter or their collaborators, maybe as gamers themselves. In fact, even in experiments without experimenter-subject communication of any kind, hypotheses are generated in the mind of the experimenter through a subjective creative process that is very poorly understood. In the field of cognitive science, they may often draw on conjectures based on the experimenter's own subjective experience of the relevant phenomena. Without this source of (almost never acknowledged) inspiration, the experiments themselves would never occur.

Many experiments also have debriefing interviews, to ensure that what subjects did in their task, e.g. while playing a video game, corresponds to what experimenters interpret they were doing. Those interviews may even be used to exclude some data from the experiment, testifying to their importance.

Taking all these considerations together, it is apparent that there is a considerable contradiction between the omnipresence of first-person experience in the cognitive science experiments and the discourse regarding its under-representation. Several authors have claimed that cognitive science is neglecting first-person experience as a rich and promising source of data. Anthony Jack and Andreas Roepstorff, for example, ask:

> What would be your response if you were told that most cognitive scientists habitually overlook a valuable source of evidence about mental processes? Every time an experiment is conducted, there are data simply waiting to be collected but persistently neglected (2002, p. 333)

From our standpoint, however, this is an overly gloomy image of what is happening in the field. Not only do scientists' first-person experiences permeate all their work, but second-person methods—any methods in which subjects deliberately communicate their first-person subjective experience to the experimenter—in fact span a very broad spectrum ranging from binary non-verbal responses to free speech. Accepting such a definition of second-person data, it is clear that cognitive science has never ceased to use them, it has simply tended to use reports of lower dimensionality and to constrain the subjects' behavior in the tasks as a way to reduce variability, allow quantification, and control for confounds. Still, when subjects intentionally and consciously choose to communicate their phenomenal experience verbally or otherwise, they are conveying part of their first-person experience and scientists trust them *prima facie*.

What we see in our review is that, overall, studies in the *Flexible* cluster tended to approach first-person experience explicitly, discussing the issues in the manuscript, while studies in the *Constrained* cluster tended to take the subject's phenomenal experience into account only implicitly. However, this does not imply that in one cluster first-person experience is included and in the other it is absent. First, since even the most rigorous and constrained psychophysics tasks rely on subjective second-person reports, throughout the vast majority of studies we examined, subjective experience was critical even when it was not explicitly acknowledged as such. Moreover, we found no sign of a well-defined frontier between methods used by those that care about first-person experience and those

who would rather avoid it entirely. Rather than discrete differences that might be implied by clustering, there was a continuum in the dimensionality of second-person reports used across studies. Despite the tendency of the dimensionality of reports to co-vary with other task constraints, this was not always the case. The 2002 study by Lutz et al. which fell under the *Constrained* cluster is one example of this. While this study was explicitly concerned with the subjects' phenomenal experience and it used open-ended second-person reports, it was in most other respects the same as more conventional studies in cognitive science in its design and methods.

3.5 The need for triangulation

Our survey also indicated that the triangulation of reports with the combination of neural and behavioral measures is underexplored. Despite the acknowledgement of its importance in the recent literature, experiments that use more than two methods to target the same phenomenon are still rare. We are aware that the criteria we used for our sampling were neither exhaustive nor unbiased, which prevents us from inferring the prevalence of certain methods with respect to others. Still, we note that behavioral measures were totally absent from 15 experiments and were very seldom rated higher than four out of ten on the scale (meaning that measures such as pupil dilation, heart rate or skin conductance were almost never used).

A more extensive inclusion of implicit behavioral or physiological third-person measures (not involving deliberate communication with the subject) in parallel with intentional second-person reports has a double advantage. On the one hand, it allows for better control of the accuracy of reports. The subjects' awareness of their ongoing experience can be considered a measuring device. Like any such instrument, this human ability can be appropriately calibrated in order to improve its reliability. One way of doing so is to confront the subjects' measurements (the reports) with results from other instruments (behavioral and physiological measures) with which we hypothesize them to be correlated. On the other hand, the use of reports may also enrich the understanding we have of the behavioral and neural measures. A good example of this can be found in the 2002 study by Lutz, Varela and colleagues, in which detailed reports about a simple subjective visual experience guided the interpretation of neural data. The training of subjects in the identification of different types of subjective experience during the task allowed for the subdivision of trials into subtle phenomenological clusters which were subsequently correlated with synchrony patterns of neural activity that would have otherwise been dismissed as noise.

3.6 The road to operationalization

There was a wide range of phenomena in the studies we reviewed. In the *Flexible* cluster, the targets of study covered a range of subjective experiences that could be considered somewhat *ineffable*: difficult to define and to communicate in a manner

that is certain to be understood similarly by subjects and experimenters. An example of this is the sense of ownership over a particular body part (Susuki et al. 2013). Such a sensation is not straightforward to explain in words, but can be demonstrated through its violation via a multisensory illusion. In the *Constrained* cluster, on the other hand, many of the phenomena being studied are relatively easy to operationalize. For example, many sensory or perceptual phenomena are tightly coupled to an external stimulus and can be readily demonstrated, even if they have a subjective component. Binocular rivalry is one such example. However, the *Constrained* cluster also included examples of seemingly ineffable phenomena such as "confidence" or "sense of agency". In a majority of these cases, the experimenters are working with concepts that have been effectively operationalized, typically through a series of studies, possibly spanning work over many years by many labs.

Given these characteristics, we suggest that the studies represented by the *Flexible* and *Constrained* clusters reflect not only distinct and complementary approaches, but are appropriate to two different stages in the development of an understanding of subjective phenomena. The *Flexible* cluster is typical of an initial stage of investigation in which the phenomena are not yet pinned down and defined in a tractable and communicable manner. The *Constrained* cluster reflects a more refined stage, in which the phenomena are (or should be in principle) more securely operationalized. When the target of study is truly difficult and novel, the less-constrained initial stage may be a necessary precursor of the more constrained stage. Less-constrained tasks and richer phenomenological reports would be critical to avoid imposing overly strong a priori assumptions on something that is not yet entirely understood. Even phenomena that seem commonplace and consensual may in fact benefit from such an exploratory scrutiny.

It is important that experimenters are cognizant of the status of the concepts that they are dealing with, as failure to do so can lead to important problems in the interpretation of what appear to be simple phenomena.[7] Terms for ineffable phenomena are also sometimes deceptively simple. For example, "happiness" is a widely used word that may cover a wide range of possible phenomenological states, ranging from physical well-being to emotional valence to pleasure or even satiation. Asking subjects to simply rate their happiness may result in their reporting a wide range of subjective phenomena. It is important that experimenters recognize that there is always a gap between what subjects are experiencing and what they are reporting. In many cases, different subjects may be experiencing different things, even if they are lumping them under the same label (i.e. lack of precision). Subjects may also, by the nature of the constraints imposed by the experimental setup, experience precisely the same thing, yet still systematically mislabel it (i.e. precision without accuracy). Unintentional experimenter bias and the tendency for subjects to unwittingly conform to the expectations of the experimenter can undermine the honest intentions of both sides to communicate about subjective experience.

---

[7] LeDoux (2014) has recently alerted us against this "substitution" fallacy (Kahneman 2011, p. 12), claiming that when we do research on fear responses in rats, we are missing confirmation of the most crucial element of fear: the subjective experience of feeling frightened.

These potential pitfalls highlight the importance and difficulty of operationalizing first-person experiences and point to the critical role of communication between the experimenter and the subject. The instructions given to a subject, which customarily receive only cursory description, require that subjects truly internalize and understand instructions. This will be difficult if not impossible to ascertain by a simple verbal affirmation, yet this is often the basis for the inclusion or exclusion of a subject's data. The simplest form of experience, such as the direction of movement in a random dot kinematogram, is something an experimenter can likely share and communicate with ease with any experimental subject. But even slightly more nuanced experiences may not be so consensual between the experimenter and subject. These considerations will be magnified when working with subjects whose backgrounds differ from the experimenter's, such as with subjects with clinical disorders.

It is our view that the sciences of the mind have always depended upon the careful articulation of first-person experience. This has been the case, despite attempts to deny it, through ever-changing viewpoints on subjective experience, from the earliest days of psychology, to introspectionism, through behaviorism and all the way to contemporary cognitive science. The epistemological conundrums that challenged introspectionism were never solved through a secure and universal procedure and they are not limited to studies that explicitly grapple with subjective experience. The overarching challenge of how to translate the unique and ineffable into the clear and reliable continues to pervade the cognitive sciences. Nevertheless, science has succeeded in providing insight into seemingly any particular subjective phenomenon that has been tackled. This suggests that, while neither wholesale dismissal nor blanket assurances are possible, individual problems are amenable to progress. A path forward to address any particular topic calls for confronting its phenomenal aspects explicitly and head on, recognizing the importance and difficulty of the process, and bringing to bear as many sources of knowledge as are available.

Compliance with ethical standards

References

Andersen, S. B., Moore, R. A., Venables, L., & Corr, P. J. (2009). Electrophysiological correlates of anxious rumination. *International Journal of Psychophysiology, 71,* 156–169.

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience, 7,* 542–548.

Baranski, J. V., & Petrusic, W. M. (1994). The calibration and resolution of confidence in perceptual judgments. *Perception and Psychophysics, 55,* 412–428.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561–571.

Beenfeldt, C. (2008). A philosophical critique of heterophenomenology. *Journal of Consciousness Studies, 15,* 5–34.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics, 29,* 336–342.

Buckner, R. L., Bandettini, P. A., O'Craven, K. M., Savoy, R. L., Petersen, S. E., Raichles, M. E., et al. (1996). Detection of cortical activation during averaged single trials of a cognitive task using functional magnetic resonance imaging. *PNAS, 93,* 14878–14883.

Buxton, W., Myers, B. A. (1986). A study in two-handed input. In *Proceedings of CHI'86* (pp. 321–326).

Carhart-Harris, R. L., Muthukumaraswamy, S., Roseman, L., Kaelen, M., Droog, W., Murphy, K., et al. (2016). Neural correlates of the LSD experience revealed by multimodal neuroimaging. *PNAS, 113,* 4853–4858.

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences, 32,* 121–182.

Caspar, E. A., Christensen, J. F., Cleeremans, A., & Haggard, P. (2016). Coercion changes the sense of agency in the human brain. *Current Biology, 26,* 1–8.

Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2012). An online neural substrate for sense of agency. *Cerebral Cortex, 23,* 1031–1037.

Charpentier, C. J., De Martino, B., Sim, A. S., Sharot, T., & Roiser, J. P. (2015). Emotion-induced loss aversion and striatal-amygdala coupling in low-anxious individuals. *Social Cognitive and Affective Neuroscience, 11,* 569–579.

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience, 20,* 115–125.

Chirimuuta, M. (2014). Psychophysical methods and the evasion of introspection. *Philosophy of Science, 81,* 914–926.

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *PNAS, 106,* 8719–8724.

Corti, K., & Gillespie, A. (2014). Revisiting Milgram's cyranoid method: Experimenting with hybrid human agents. *The Journal of Social Psychology.* https://doi.org/10.1080/00224545.2014.959885.

Cosmelli, D., David, O., Lachaux, J.-P., Martinerie, J., Garnero, L., Renault, B., et al. (2004). Waves of consciousness: Ongoing cortical patterns during binocular rivalry. *NeuroImage, 23,* 128–140.

Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology. The collected works of Mihaly Csikszentmihalyi.* Dordrecht: Springer.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts.* New York: Viking Penguin.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79,* 1–37.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5,* 303–305.

Dennett, D. C. (1991). *Consciousness explained.* New York: Little, Brown and Company.

Dennett, D. C. (1996). *The intentional stance.* Cambridge, MA: The MIT Press.

Dennett, D. C. (2007). Heterophenomenology reconsidered. *Phenomenology and Cognitive Sciences, 6,* 247–270.

Ericsson, K. A. (2003). Valid and non-reactive verbalization of thoughts during performance of tasks. *Journal of Consciousness Studies, 10,* 1–18.

Erowid experience vaults. https://www.erowid.org/. Accessed 2 Apr 2017.

Eslinger, P. J., & Damasio, A. R. (1986). Preserved motor learning in Alzheimer's disease: Implications for anatomy and behavior. *The Journal of Neuroscience, 3,* 3006–3009.

Evans, G. W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 13–24.

Feest, U. (2012). Introspection as a method and introspection as a feature of consciousness. *Inquiry, 55,* 1–16.

Feinstein, J. S., Adolphs, R., Damásio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current Biology, 21,* 34–38.

Fleming, S. M., & Dolan, R. L. (2009). Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Consciousness and Cognition, 19,* 352–363.

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *The Journal of Neuroscience, 32,* 6117–6125.

Fontani, G., Corradeschi, F., Felici, A., Alfatti, F., Migliorini, S., & Lodi, L. (2005). Cognitive and physiological effects of Omega-3 polyunsaturated fatty acid supplementation in healthy subjects. *European Journal of Clinical Investigation, 35,* 691–699.

Fox, M. C., Ericsson, K. A., & Best, R. (2010). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137,* 316–344.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137,* 316–344.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron, 69,* 548–562.

Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences, 3,* 105–114.

Gallagher, S. (2002). Experimenting with introspection. *Trends in Cognitive Sciences, 6,* 347–348.

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *NeuroImage, 16,* 814–821.

Gallagher, S., & Overgaard, M. (2006). Introspections without introspeculations. In M. Aydede (Ed.), *Pain. New essays on its nature and the methodology of its study* (pp. 277–289). Cambridge, MA: MIT Press.

Gallagher, S., Reinerman-Jones, L., Janz, B., Bockelman, P., & Trempler, J. (2015). *A neurophenomenology of awe and wonder. Towards a non-reductionist cognitive science.* New York: Palgrave Macmillan.

Garrison, K. A., Scheinost, D., Worhunsky, P. D., Elwafi, H. M., Thornhill, T. A., IV, Thompson, E., et al. (2013). Real-time fMRI links subjective experience with brain activity during focused attention. *NeuroImage, 81,* 110–118.

Goldman, A. I. (2004). Epistemology and the evidential status of introspective reports. *Journal of Consciousness Studies, 11,* 1–16.

Gould van Praag, C. D., Garfinkel, S., Ward, J., & Bor, D. (2016). Automaticity and localization of concurrents predicts colour area activity in grapheme-colour synaesthesia. *Neuropsychologia, 88,* 5–14.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273,* 1699–1702.

Griffiths, R. R., Richards, W. A., McCann, U., & Jesse, R. (2006). Psilocybin can occasion mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology (Berl), 187,* 268–283.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5,* 382–385.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron, 37,* 877–888.

Hatfield, G. (2005). Introspective evidence in psychology. In P. Achinstein (Ed.), *Scientific evidence: Philosophical theories and applications* (pp. 259–286). Baltimore, London: Johns Hopkins University Press.

Hoyer, W. D., & Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *The Journal of Consumer Research, 17,* 141–148.

Hurlburt, R. T. (1997). Randomly sampling thinking in the natural environment. *Journal of Consulting and Clinical Psychology, 65,* 941–949.

Hurlburt, R. T., & Heavey, C. L. (2006a). *Exploring inner experience. The descriptive experience sampling method.* Amsterdam, Philadelphia: John Benjamins Publishing Company.

Hurlburt, R. T., & Heavey, C. L. (2006b). *Exploring inner experience. The descriptive experience sampling method.* Philadelphia: John Benjamins Publishing Company.

Jack, A. I., & Roepstorff, A. (2002). Introspection and cognitive brain mapping: From stimulus-response to script-report. *Trends in Cognitive Sciences, 6,* 334–338.

Jack, A. I., Roepstorff, A. (eds.). (2003–4). *Trusting the subject? The use of introspective evidence in cognitive science* (Vol. 1, 2). Thorverton: Imprint Academic.

James, W. (1890). *The principles of psychology*. New York: Holt.

Jo, H.-G., Wittman, M., Borghardt, T. L., Hinterberger, T., & Schmidt, S. (2014). First-person approaches in neuroscience of consciousness: Brain dynamics correlate with the intention to act. *Consciousness and Cognition, 26,* 105–116.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kempf, D. S. (1999). Attitude formation from product trial: Distinct roles of cognition and affect for hedonic and functional products. *Psychology & Marketing, 16,* 35–50.

Koepp, M. J., Gunn, R. N., Lawrence, A. D., Cunningham, V. J., Dagher, A., Jones, T., et al. (1998). Evidence for striatal dopamine release during a video game. *Nature, 393,* 266–268.

Kühn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex, 49,* 1935–1942.

Lahlou, S., Le Bellu, S., & Boesen-Mariani, S. (2015). Subjective evidence based ethnography: Method and applications. *Integrative Psychological and Behavioral Science, 49,* 216–238.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133,* 339–354.

LeDoux, J. E. (2014). Coming to terms with fear. *PNAS, 111,* 2871–2878.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). *Brain, 106,* 623–642.

Lutz, A., Brefczynski-Lewis, J., Johnstone, T., & Davidson, R. J. (2008). Regulation of the neural circuitry of emotion by compassion meditation: Effects of meditative expertise. *PLoS ONE, 3,* e1897.

Lutz, A., Greischar, L. L., Rawlings, N. B., Ricard, M., & Davidson, R. J. (2004). Long-term meditators self-induce high-amplitude gamma synchrony, during mental practice. *PNAS, 101,* 16369–16373.

Lutz, A., Lachaux, J.-P., Martinerie, J., & Varela, F. J. (2002a). Guiding the study of brain dynamics by using first person data: Synchrony patterns correlate with ongoing conscious states during a simple visual task. *PNAS, 99,* 1586–1591.

Lutz, A., et al. (2002b). Guiding the study of brain dynamics by using first person data: Synchrony patterns correlate with ongoing conscious states during a simple visual task. *PNAS, 99,* 1586–1591.

Marcel, A. I. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15,* 197–237.

Marks, D. F. (1973a). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64,* 7–24.

Marks, D. F. (1973b). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64,* 17–24.

Metcalfe, J., & Greene, M. J. (2007). Metacognition of Agency. *Journal of Experimental Psychology: General, 136,* 184–199.

Michael, J., Sandberg, K., Skewes, J., Wolf, T., Blicher, J., Overgaard, M., et al. (2014). Continuous theta-burst stimulation demonstrates a causal role of premotor homunculus in action understanding. *Psychological Science, 25,* 963–972.

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

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111,* 275–279.

Osu, R., Hirai, S., Yoshioka, T., & Kawato, M. (2004). Random presentation enables subjects to adapt to two opposing forces on the hand. *Nature Neuroscience, 7,* 111–112.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsoy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition, 15,* 700–708.

Parnas, J., & Henriksen, M. G. (2014). Disordered self in the schizophrenia spectrum: A clinical and research perspective. *Harvard Review of Psychiatry, 22,* 251–265.

Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292,* 510–512.

Penfield, W. (1958). Some mechanisms of consciousness discovered during electrical stimulation of the brain. *Proceedings of the National Academy of Sciences, 44,* 51–66.

Penfield, W. (1959). The interpretive cortex. *Science, 129,* 1719–1725.

Persaud, N., McLeod, P., & Cowey, A. (2007a). Post-decision wagering objectively measures awareness. *Nature Neuroscience, 10,* 257–261.

Persaud, N., McLeod, P., & Cowey, A. (2007b). Post-decision wagering objectively measures awareness. *Nature Neuroscience, 10,* 257–261.

Petitmengin, C. (2006). Describing one's subjective experience in the second person: As interview method for the science of consciousness. *Phenomenology and Cognitive Science, 5,* 229–269.

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.

Petitmengin-Peugeot, C. (1999). The intuitive experience. *Journal of Consciousness Studies, 6,* 43–77.

Phelps, E. A., O'Connor, K. J., Gatenby, C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience, 4,* 437–441.

Piccinini, G. (2009). First-person data, publicity and self-measurement. *Philosophers' Imprint, 9,* 1–6.

Piccinini, G. (2010). How to improve on heterophenomenology. The self-measurement methodology of first-person data. *Journal of Consciousness Studies, 17,* 84–106.

Preuschoff, K., t Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience, 5,* 115.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia—A window into perception, thought and language. *Journal of Consciousness Studies, 8,* 3–34.

Ramsoy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences, 3,* 1–23.

Ratcliff, R., & Rouder, J. N. (1998). Modelling response times for two-choice decisions. *Psychological Science, 9,* 347–356.

Roberts, S. (2004). Self-experimentation as a source of new ideas: Ten examples about sleep, mood, health, and weight. *Behavioral and Brain Sciences, 27,* 227–288.

Roepstorff, A. (2001). Brains in scanners: An Umwelt of cognitive neuroscience. *Semiotica, 134,* 747–765.

Rosenthal, R., & Rosnow, R. L. (2009). *Artifacts in behavioral research: Robert Rosenthal and Ralph L. Rosnow's classic books.* New York: Oxford University Press.

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a statistical computation in the human sense of confidence. *Neuron, 90,* 499–506.

Schlegel, A., Alexander, P., Sinnott-Armstrong, W., Roskies, A., Tse, P. U., & Wheatley, T. (2015). Hypnotizing Libet: Readiness potentials with non-conscious volition. *Consciousness and Cognition, 33,* 196–203.

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

Sidarus, N., & Haggard, P. (2016). Difficult action decisions reduce the sense of agency: A study using the Eriksen flanker task. *Acta Psychologica, 166,* 1–11.

Simon, H. A. (1980). Cognitive science: The newest science of the artificial. *Cognitive Science, 4,* 33–46.

Slagter, H. A., Lutz, A., Greischar, L. L., Nieuwenhuis, S., & Davidson, R. J. (2008). Theta phase synchrony and conscious target perception: Impact of intensive mental training. *Journal of Cognitive Neuroscience, 21,* 1536–1549.

Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 544–551.

Smith, R. E., & Swinyard, W. R. (1983). Attitude-behavior consistency: The impact of product trial versus advertising. *Journal of Marketing Research, 20,* 257–267.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654.

Susuki, K., Garfinkel, S. N., Critchley, H. D., & Seth, A. K. (2013). Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion. *Neuropsychologia, 51,* 2909–2917.

Titchener, E. (1901a). *Experimental psychology. A manual of laboratory practice.* London: MacMillan.

Titchener, E. (1901b). *Experimental psychology. A manual of laboratory practice, vol. I, qualitative experiments. Part I, student's manual.* London: MacMillan.

Trull, T. J., Solhan, M. B., Tragesser, S. L., Jahng, S., Wood, P. K., Piasecki, T. M., et al. (2008). Affective instability: Measuring a core feature of borderline personality disorder with ecological momentary assessment. *Journal of Abnormal Psychology, 117,* 647–661.

Varela, F. (1996). Neurophenomenology: A methodological remedy to the hard problem. *Journal of Consciousness Studies, 3,* 330–350.

Varela, F. J., & Shear, J. (1999a). First-person methodologies: What, why, how? *Journal of Consciousness Studies, 6,* 1–14.

Varela, F. J., & Shear, J. (Eds.). (1999b). *The view from within: First-person approaches to the study of consciousness*. Bowling Green, OH: Imprint Academic.

Velmans, M. (1999). Intersubjective science. In F. J. Varela & J. Shear (Eds.), *The view from within: First-person approaches to the study of consciousness* (pp. 299–306). Bowling Green, OH: Imprint Academic.

Velmans, M. (2007). Heterophenomenology versus critical phenomenology. *Phenomenology and the Cognitive Sciences, 6,* 221–230.

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

Voss, U., Holzmann, R., Hobson, A., Paulus, W., Koppehele-Gossel, J., Klimke, A., et al. (2014). Induction of self awareness in dreams through frontal low current stimulation of gamma activity. *Nature Neuroscience, 17,* 810–812.

Ward, J., Jonas, C., Dienes, Z., & Seth, A. (2010). Grapheme-colour synaesthesia improves detection of embedded shapes, but without pre-attentive 'pop-out' of synaesthetic colour. *Proceedings of the Royal Society B, 277,* 1021–1026.

Wegner, D., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist, 54,* 480–492.

Wilson, T. D. (2003). Knowing when to ask. Introspection and the adaptive unconscious. *Journal of Consciousness Studies, 10,* 131–140.

Woodworth, R. S. (1921). *Psychology*. New York: Holt.

Woodworth, R. S. (1934). *Psychology* (3rd ed.). New York: Holt.