Converging on an understanding of the déjà vu experience.

Courtney B Aitken – University of St Andrews, Akira R O’Connor – University of St Andrews

“For a few precarious seconds, the chaplain tingled with a weird occult sensation of having experienced the identical situation in some prior time or existence… déjà vu. The subtle, recurring confusion between illusion and reality that was characteristic of paramnesia fascinated the chaplain.” (Heller, 1955, p. 235). The chaplain’s experience in Catch 22 provides a vivid illustration of a memory quirk that has fascinated creatives for centuries. The scientific study of déjà vu arrived with Boirac’s (1876) description and Arnaud’s single case study (Arnaud, 1896; see Bertrand, Martinon, Souchay, & Moulin, 2017 for a translation into English) but by then a significant body of literature describing the experience, from Saint Augustine (400AD, XII.xv.24), to Dickens (1850) and Hawthorne (1863), had already amassed.

Fundamental to references to déjà vu in popular literature, extended today into song, cinema and online media, is the phenomenologically uncanny, highly idiosyncratic nature of the experience. Without any unique observable markers, the only way those experiencing it can understand that déjà vu is not unique to them, is to describe it in all its subjectivity. In turn, those observing the occurrence of déjà vu rely on these descriptions and their own interpretations of them being accurate. The processes of introspecting and communicating this information to one another is not unique to déjà vu—it could be argued that communicating the occurrence of any cognitive process shares the same challenges—but what makes déjà vu unique (alongside related memory quirks, such as jamais vu and presque vu) is that it cannot be identified solely by its behavioural outcomes. Here, we will describe key features of the déjà vu experience, ultimately concluding that subjective experience lies at its core, in spite of disagreement about its behavioural consequences. We will then outline how three strands of déjà vu research have sought to investigate the experience, which each bring with them attendant challenges. Finally, we propose a single, convergent, prospective approach by which these collective challenges can be overcome.

Déjà vu as a Subjective Experience

Definitions of déjà vu have either concentrated solely on the feeling of uncanny familiarity, or introduced an additional conflict with an awareness that the familiarity is incorrect. Brown (2004) provided a table of 53 definitions from the scientific literature, within which the majority (40) focused solely on the feeling of familiarity, typically noting that déjà vu was experienced in the context of a situation that was novel (see Table 1). The familiarity in
question was typically defined as being juxtaposed with situational novelty, though awareness of novelty did not form a part of the definition. A smaller number also noted not just that the situation was novel, but that the awareness of this novelty was integral to the experience. These definitions (11) focused on familiarity-novelty conflict, implying that the experient’s awareness is more ‘objectively’ compelling than the feeling of familiarity, but nonetheless not requiring any specific resolution of conflict (either in favour of novelty or familiarity) for the experience to be considered déjà vu. Two definitions, however, did go beyond simply noting the internal conflict, and included its resolution in favour of novelty in the definition. Here we briefly argue that, whilst we believe the resolution in favour of the evaluation of novelty is integral to the experience, all definitions rightly prioritise the subjective nature of the experience. This in turn necessitates the methodological rigour we argue as essential for the investigation of déjà vu, and related subjective phenomena.

Table 1: Classification of Déjà vu Definitions from Brown (2004)

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarity</td>
<td>“a person comes across a place or situation that seems familiar, but which the person has never encountered before.” (Glass, Holyoak &amp; Santa, 1979, p. 63).</td>
<td>40</td>
<td>75</td>
</tr>
<tr>
<td>Familiarity-novelty conflict</td>
<td>“The experience of feeling ‘I have been here before’ or ‘I have lived through this before’, together with intellectual awareness that this is not so” (McKellar, 1957, p. 200).</td>
<td>11</td>
<td>21</td>
</tr>
<tr>
<td>Familiarity-novelty conflict resolved in favour of novelty</td>
<td>“… subject feels strongly that he has seen or done something before, but usually remains logically convinced that he has not” (Pickford, 1940, p. 152).</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

A related experience that stresses the definitional importance of memory conflict resolution in standard déjà vu, is déjà vécu. Déjà vécu experients describe similar feelings to those experiencing déjà vu—familiarity alongside awareness of novelty—but end up acting in ways very different to them. The divergence is primarily that déjà vécu experients resolve their memory conflict in favour of their erroneous familiarity, withdrawing from the activity associated with the sensation and even confabulating other details to counterfactually explain how they could actually have experienced the situation before (e.g. Moulin, Conway, Thompson, James, & Jones, 2005). Patient AKP, a déjà vécu case-study, was remarkable not because he experienced sensations described initially as déjà vu, but because his feeling of recognition was so compelling that he lived as though it was correct (Moulin et al.,
2005; notably, this was also true for the first déjà vu case study, Arnaud’s, 1896, patient Louis). AKP stopped reading the newspaper and watching television because each time he thought to, he became convinced he had already done so—he also confabulated justifications for these withdrawal behaviours, such as going to the newsagent early in the morning and reading the newspaper there. Moulin and colleagues noted the behavioural change driven by AKP’s illusory memory sensations, identified it as being qualitatively different to déjà vu in its behavioural outcome, and referred to AKP’s experience as déjà vécu (literally “already lived”). If déjà vu were characterised by erroneous familiarity or a conflict between familiarity and an awareness that this familiarity conflicted with other knowledge the experient held about their situation, AKP’s situation would not be so striking. It would not be unusual for déjà vu to drive behaviour, leading to false recognition, at least some of the time. Were this the case, those experiencing déjà vu would be more likely to show observable behaviours of the kind that brought AKP to the attention of his doctors (albeit less frequently). What is apparent from our responses to the unusual behaviours of those with déjà vécu is we have an intuitive understanding that déjà vu experiencers range in their evaluations from ambivalence about novelty and familiarity (parallel sensations potentially supported by neuroanatomically distinct systems e.g. Kafkas & Montaldi, 2018; Duszkiewicz, McNamara, Takeuchi & Genzel, 2019) to certainty that their familiarity is false. Given this understanding of the experience, it follows that memory conflict resolution in favour of novelty awareness plays an important role in déjà vu experience, whilst resolution in favour of familiarity is indicative of different memory phenomena ranging from false memory to delusion.

Whilst we argue that the resolution in favour of novelty is necessary during déjà vu, it is not sufficient—déjà vu cannot simply be defined by its behavioural outcome (or lack thereof). Correct rejection of novel stimuli (novel stimuli correctly identified as “new”) far more typically results from everyday memory evaluations not at all linked to déjà vu. It should be noted that this is the case whether one agrees with our position on memory conflict resolution in déjà vu, or takes a broader view of déjà experiences in which they can be experienced alongside correct rejections and false alarms (novel stimuli incorrectly identified as “old”). False alarms also occur somewhat frequently in everyday life and are ubiquitous in standard cognitive tests of memory. What is required to define déjà vu therefore is a combination of the objective outcome (correct identification of a novel stimulus as novel) alongside a subjective state (awareness that familiarity and an evaluation of novelty have come into conflict). This latter insight into the phenomenological experience underpinning the outcome can only be gained from introspection, and despite the attempts of scientists to define the experience as objectively as possible, it is notable that words such as strange (six occurrences), illusion
(five occurrences), uncanny and weird (two occurrences each) appear in Brown’s table of definitions. Descriptions of déjà vu inevitably nod towards the uncanny, and rely on the experient to select those experiences that satisfy this nebulous criterion and therefore qualify as déjà vu, distinct from those experiences that do not.

We now turn to a range of methods by which déjà vu has been researched. As we do so, it will be apparent that some lend themselves to the exploration of the subjective component of the experience more than others. We will revisit this issue more critically later, with particular emphasis on the malleability of memory and the correspondence between déjà vu experienced by clinical and experimental samples, and the naturalistic déjà vu experience that drove Heller and Dickens to revel in its curious nature.

**Priorities in Déjà vu Research**

The study of déjà vu has taken three approaches driven by the priorities of researchers. The first approach prioritises study of the naturalistic experience, taking broad samples, representative of the neurotypical population, but limiting the depth with which the nature of the experience can be explored (reviewed extensively in Brown, 2003). The second prioritises study of those with the most frequent experience. This allows déjà vu to be studied more deeply through the participant’s increased likelihood of being able to bring recent relevant experiences to mind, or the increased likelihood of occurrence in the presence of the researcher. Given the infrequency of déjà vu (relative frequency estimates in broad samples suggest 1 to 3 times per year; Brown, 2004) this necessarily involves the study of special samples who either have medical conditions or circumstances that increase these individuals’ likelihood of experiencing déjà vu (e.g. Labate et al., 2011). The third approach combines the previous two, prioritising the in-depth study of déjà vu in samples of neurotypical adults (e.g. Urquhart & O’Connor, 2014). This combination necessarily leads to compromise in the extent to which the infrequently occurring experience can be studied in the lab. As such, this third approach has primarily been enabled by the growing interest in experimentally-generated analogues of déjà vu.

**Naturalistic Experience in Broad Samples.** The approach which has yielded the most enduring findings about déjà vu is the retrospective questionnaire. Often with little context, large number of participants answer a question (or questions) relating to déjà vu, typically alongside other items relating to their demographics, habits and circumstances of the experience. Researchers take an associative analytic approach to find relationships in the data (e.g. Chapman & Mensh, 1951; Harper, 1969; Neppe, 1983; Richardson & Winokur,
1967; Gallagher, Kumar, & Pekala, 1994; Silberman, Sussman, Skillings, & Callanan, 1994; Wiener, 1992; Kohr, 1980; Palmer, 1979), with a view to generalising from the representative sample to the neurotypical population. Different surveys take opposing positions on identifying the experience in question as ‘déjà vu’, with only very few choosing to do so (e.g. Gallup & Newport, 1991), and most providing a definition of the experience without using the words ‘déjà vu’ themselves (e.g. O’Connor & Moulin, 2013; Neppe, 1983). If the characteristics of déjà vu are examined, different approaches are also taken to their measurement, with some asking participants about typical experiences (e.g. Brown, Porter & Nix, 1994), and others asking participants to think back to a specific experience, such as the last déjà vu they remember (e.g. Web and Texas Surveys, cited in Brown, 2004).

Observable Experience in Special Samples. Clinical and unusual samples of participants who experience déjà vu more frequently than the general population are interrogated, typically in more detail than the samples above, on the basis that their experiences are more frequent and, in some situations, more likely to be reproduced in the laboratory. A case-study approach has been taken with patients and healthy people experiencing noteworthy (e.g. Taiminen & Jaaskelainen, 2001; Singh, 2007) and mundane instances of déjà vu (e.g. O’Connor & Moulin, 2008), but the majority of this type of research focuses on special samples that report déjà vu at higher rates than are typically observed in the general population. These participants may exhibit higher reporting rates because this reflects their natural tendency (e.g. Brázdil et al. 2012), or because they have a particular interest in the experience and have chosen to be particularly attentive to its occurrence. For example, Leeds (1944) prospectively kept a diary of his déjà vu experiences, taking advantage of his commitment to recording these episodes, and his interest in the area, to ensure that his records were made near their time of occurrence. This is an important point as regards the distinction between the broad samples inquiry described above, and the special samples inquiry here. The proximity to researchers in the current approach provides greater opportunity by which the minutiae of the experience can be interrogated, giving greater insight into the experience being investigated.

A large proportion of special samples studies have been carried out on people with temporal lobe epilepsy (e.g. Bancaud, Brunet-Bourgin, Chauvel, & Halgren, 1994) both because they often experience déjà vu as part of their pre-seizure aura, and because they can require surgical intervention, with the opportunity to observe the antecedents and experiences of déjà vu-like phenomena in the laboratory (Illman, Butler, Souchay, & Moulin, 2012). For example, Penfield analysed verbal accounts given as epileptic patients underwent cortical stimulation and experienced ‘dreamy states’ (déjà vu-like experiences induced by cortical...
brain stimulation; e.g. Penfield & Perot, 1963). The historical association of déjà vu with temporal lobe epilepsy in this approach has been pivotal in shaping our understanding of the phenomenological experience, the neural activity that might underpin déjà vu and whether clinical experiences are equivalent to apparently benign déjà vu experiences. It is for this reason, in our later discussion of observable experiences in special samples, we focus on déjà vu research involving temporal lobe epilepsy.

**Experimental Approaches in Broad Samples.** The approaches above provide the non-overlapping advantages of either dealing with a broad sample one would expect to most readily generalise to the general population, or dealing with an observable experience that can be studied in the laboratory. There has been growth in attempts to marry the two, by studying experimental analogues of déjà vu in the general population. A number of experimental studies have attempted to engineer situations in which participants’ cognitive states most closely match those which the experimenters determine correspond to the cognitive state during déjà vu. A notable example of a coherent programme of work in this domain is from Cleary (e.g. Cleary & Reyes, 2009; Cleary et al., 2012), who has generated contextual similarity between experimental situations, resulting in neurotypical participants reporting déjà vu at high rates in the laboratory. Similarly, work from our group (e.g. Urquhart & O’Connor, 2014; Urquhart, Sivakumaran, Macfarlane, & O’Connor, 2018) has developed to emphasise the resolved conflict that we view as important to déjà vu and used the high rates of resultant déjà vu to attempt to identify the neural correlates of the experimental analogue using functional magnetic resonance imaging (fMRI).

There is, of course, plenty of work which is not so straightforward to classify. Banister and Zangwill (1941) used hypnosis to generate an analogue of déjà vu, a procedure replicated in hypnotic ‘virtuosos’ by O’Connor, Barnier and Cox (2008). This illustrates a potentially undesirable combination of an experimental (non-naturalistic) analogue observable in a highly exclusive sub-sample of the population (only 10% satisfy the highly hypnotisable “virtuoso” criteria). Nonetheless, the diversity of approaches, and the difficulty in converging on a single approach rightly suggests that each runs into certain difficulties absent in the others. We now detail the issues that affect each category of approach, some which are well established, and others which are only now coming to prominence.

**Methodological Challenges**

**Naturalistic Experience in Broad Samples: The Self-Report Questionnaire.** There is considerable heterogeneity in the self-report formats used to solicit déjà vu reports.
Nevertheless, since its publication in 1994, many studies have based their assessment items on the Inventory for Déjà vu Experiences Assessment questionnaire (IDEA; Sno, Schalken, Onghe, & Koeter, 1994). The items included in the IDEA have been used extensively to establish differences (or lack thereof) between groups (including clinical groups, e.g. Adachi et al., 2007), has opened up the exploration of the non-clinical experience as a research priority in its own right (e.g. Kusumi, 1996, 1998), and has been adapted for use within neuroimaging work (e.g. Brázdil et al., 2012). The widespread use of a reliable retrospective measure focusing on the qualities of the déjà vu experience (e.g. intensity, duration, effects on the experient, modalities the experience occurred in) has raised the profile of the subjective experience of déjà vu in the literature. However, given the issues we discuss below relating to memory and survey design, whether retrospective self-report is a valid and reliable method overall remains to be seen.

Young adult samples—the group that experience déjà vu most frequently—would only be expected to report déjà vu three to four times per year. This inevitably leads to a lag, typically of more than a few weeks, between the experience and its report in a survey. Déjà vu is fleeting and infrequent, we have very few opportunities and a limited timeframe within which to examine our déjà vu experiences (once it has happened, we cannot evaluate it directly any longer than its fleeting occurrence, beyond which we must rely on our memories of the experience). We also know from general memory and self-report research that our memories of these types of experiences are prone to reconstruction and bias (Wagenaar, 1986). Whilst there is little research on how these factors affect déjà vu research directly, the reliability of the self-report methodology has been examined in other cognitive and behavioural domains. For example, flashbulb memories (memories for unexpected, emotional events e.g. terrorist attacks) are known to be extremely susceptible to reporting inconsistencies over time despite the experiencers’ increased confidence in the veracity of their reports (Talarico & Rubin, 2003). In false memory and eyewitness testimony research (e.g. Loftus & Palmer, 1974), asking participants leading questions on a certain event can result in memory distortion and inaccurate reporting. It could be argued that déjà vu is different from reporting flashbulb or eyewitness experiences as it is a uniquely personal event and, perhaps, less likely to be socially reconstructed in the same way flashbulb memories and eyewitness accounts are. The fleeting nature of déjà vu could be better thought of as an individual subjective experience not dissimilar to a medical symptom.

Self-report reliability research typically involves i) the assessment of discrepancies between self-reported events and corresponding records in databases; or ii) investigating the strategies people use to make judgements on the frequency and intensity of frequent,
generic behaviours (e.g. visiting the doctor, buying ice cream) and extrapolating from these to judgements of the experience in question. When people identify the frequency of an experience (like déjà vu), the strategies used to make that judgment are highly unreliable. They can draw on a range of strategies including recalling and counting individual episodes, recalling individual episodes and using these as the basis for extrapolation, using subjective sources of self-knowledge to estimate occurrences, and using general impressions of frequency in others to estimate occurrences in oneself (Brown, Williams, Barker, & Galambos, 2007; Conrad, Brown, & Cashman, 1998). If the experience is more subjective (e.g. related to emotion), estimates of those behaviours tend to be based on self-knowledge or general impressions rather than any relevant experience (Brown et al., 2007; Conrad et al., 1998). It is difficult to obtain a clear picture of how these strategies influence the whole retrospective report but what is apparent is that reliance on strategies like “recall and count” results in underreporting of frequency estimates (Lugtig, Glasner, & Boeve, 2016). The wider the interval participants are asked to consider the event within (e.g. over the past year or within a year), the higher rates of underreporting (Lugtig et al., 2016). For déjà vu experiences, which are thought to be temporally distant, asking participants to recount how many times in the past year results in underreporting of the experience through retrieval failure or forgetting of individual experiences (Lugtig et al., 2016; Schwarz, 2007a).

A related line of research we consider when evaluating the reliability of retrospective reports comes from clinical research investigating the reporting of generic symptoms (e.g. pain, fatigue, mood; like déjà vu, generic symptoms are experiential and usually temporary). The design of symptom questionnaires is comparable to déjà vu questionnaires in that they require participants to give indiscriminate ratings (usually of symptom frequency and intensity) aggregated across varying reporting periods (Van den Bergh & Walentynowicz, 2016; Zijlema et al., 2013). Assessing accuracy and bias in symptom reporting more often combines prospective reporting of symptoms (i.e. reporting the experience when it occurs) or ecological momentary assessments (EMA; repeated sampling of current symptoms and behaviours) with corresponding retrospective symptom reports given at a later date (Conner & Barrett, 2012). Typically, retrospective symptom intensity reports are inflated compared to their prospective reports or averaged EMAs (Walentynowicz, Van Diest, Raes, & Van den Bergh, 2017). As with survey methodology research, retention intervals are influential in impacting the reporting of generic symptoms. For generic and medically unexplained symptoms, longer retention periods result in less consistency between the experience and retrospective report (Broderick et al., 2008; Houtveen & Oei, 2007; Robinson & Clore, 2002). Broderick et al. (2008) obtained several momentary pain and fatigue ratings a day from patients in a rheumatology practice for a month-long period. The averaged momentary
ratings were compared to retrospective ratings made over varying reference periods from 1 day up to 28 days. Despite increasing difficulty of recall with increased retention intervals, there was substantial correspondence between reports over the recall periods, but retrospective ratings of symptoms were consistently over-reported compared to averaged momentary ratings. Given the correspondence of reports over short and long retention intervals, Broderick et al. (2008) suggested patients use beliefs about pain and fatigue rather than episodic recall of their experiences to inform their reports.

Models of concurrent and retrospective reports of emotion (e.g. Robinson & Clore, 2002) would suggest that experience reports at the time of the event reflects experiential knowledge (and as such is the closest insight to the experience that we can obtain from participants). Retrospective reports, on the other hand, are more complex because we have to recognise that over time there is a loss of episodic detail and so rather than the report being based on an accurate memory of the experience, it may be based on semantic or situation specific knowledge available to the participant at the time of reporting (Houtven & Oei, 2007). When participants make several reports of symptoms in real time, greater variability of those symptoms over the reporting period results in higher symptom levels recalled relative to the averaged prospective reports (Stone, Schwartz, Broderick, & Shiffman, 2005). Other memory biases (e.g. peak and recency effects) reinforce the finding that recalled ratings of previous symptom experiences are not a simple average of the experience. In peak and recency effects (particularly acute in pain experiences), the most intense part of the experience and the end of the experience disproportionately influence the overall retrospective estimation (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Stone et al., 2005; Walentynowicz, Bogaerts, Van Diest, Raes, & Van den Bergh, 2015).

Symptom levels at reporting (i.e. higher symptoms at recall related to greater overreporting), individual differences in affective traits (e.g. anxiety, higher anxiety related to greater overreporting) and the affective state of the experient at the time of event (e.g. higher state anxiety related to greater overreporting) are the three further factors which affect the accuracy of retrospective symptom reports (Bogaerts et al., 2015; Broderick et al., 2008; Gedney, Logan, & Baron, 2003; Howren & Suls, 2011; Petersen, Van Staeyen, Vögele, von Leupoldt, & Van den Bergh, 2015; Van den Bergh & Walentynowicz, 2016). People with clinical anxiety are a cohort known to report experiencing déjà vu more frequently and more intensely during periods of high anxiety (Harper & Roth, 1962; Wells et al., 2014; Wells, O’Connor, & Moulin, 2018). When we ask participants about covariates of their experience (e.g. stress or anxiety), we are asking them to make judgements on the frequency of those experiences and the contexts in which they occur. However, concurrent and retrospective
reporting studies on generic somatic experiences suggest that people typically have very limited access to relevant episodic details that would inform an accurate response (Schwarz, 2007b). With all this in mind, it becomes very difficult to ensure that differences in reporting rates between people with anxiety and non-clinical controls are genuine differences and not driven by individual differences in symptom (over-)reporting (Schwarz, 2007b; Van den Bergh & Walentynowicz, 2016). A similar issue arises when asking participants about covariates risks reconstructing their experience or judgments in retrospective reports. For example, a participant might respond positively to a question asking if anxiety is a déjà vu trigger not because they have recalled instances of experiencing déjà vu while anxious but because the question has reconstructed this experience, or it confirms a naïve theory of covariation.

**Observable Experience in Special Samples: Experiential Equivalence.** The prevalence of déjà vu as a component of the pre-seizure aura has linked the study of déjà vu to epilepsy research. Recently though, there has been growing uncertainty regarding the subjective equivalence of the ictal (epileptic) experience of déjà vu, and experience within the broader population. Fukao, Murai, Yamada, Sengoku, & Kusumi (2005) used a retrospective questionnaire to examine phenomenological differences in déjà vu and jamais vu (i.e. experiencing a sense of unfamiliarity for a situation that should feel familiar) experiences between people with epilepsy and non-clinical controls. They found differences in the situations that were associated with experiencing déjà vu in both samples. Non-clinical participants associated being in new places and being in unfamiliar situations with experiencing déjà vu while participants with epilepsy associated being in familiar situations with having the experience. For both groups of participants, déjà vu and jamais vu only lasted several to tens of seconds but evoked different feelings in participants (e.g. déjà vu evoked surprise in nonclinical participants but not in participants with epilepsy). Similarly, Adachi et al. (2010) found déjà vu with seizure recognition was more often associated with psychopathological characteristics such as experiencing dissociation and unpleasant feelings (e.g. fear, discomfort) while non-clinical controls included in the study did not associate déjà vu with other psychopathologies. Warren-Gash & Zeman (2014) reported similar phenomenological differences between non clinical déjà vu and experiences associated with TLE. In the end concluding that epileptic déjà vu on its own could potentially be identified by concurrent symptoms and experiences associated with epilepsy such as fatigue and headaches.

Given the possibility of distinguishable phenomenological and psychological experiences between déjà vu experiences in TLE and those in apparently non-clinical experiences, there
has also been a drive to understand whether “normal” déjà vu experiences in the non-epileptic population are underpinned by ictal events. Labate et al. (2011) have argued that déjà vu experiences in “healthy” participants could be indicative of benign TLE where individuals present with ictal symptoms including prominent déjà vu experiences but do not have any febrile seizures. They speculated that since déjà vu is sometimes the only symptom of a seizure in patients of benign TLE, it could mean they fall somewhere on the TLE spectrum. In a commentary article on Brázdil et al. (2012), and in response to their argument that “physiological” déjà vu in healthy participants could be likened to a small seizure, Labate and Gambardella (2013) argued that some of their apparently healthy control participants could be people with benign TLE. Given 40% of those diagnosed with benign TLE show MRI evidence of hippocampal sclerosis associated with drug-resistant epilepsy, the authors in their commentary suggested an MRI protocol to help visualize the mesial temporal structures in these participants and rule out the possibility of déjà vu as part of benign TLE. Brázdil & Zeman (2013) argued that resolving this issue of categorizing ambiguous déjà vu experiences requires having a more robust diagnosis process for mild cases of TLE more generally. The experience of déjà vu within (benign) TLE is nebulous even with the use of MRI protocols for diagnosis (60% of those diagnosed with benign TLE show no hippocampal sclerosis) making it extremely difficult to rule out “healthy” controls as having benign TLE.

Beyond this debate, Perucca et al. (2017) assessed whether genetic factors (established through family histories) may be contributing to a diagnosis of benign TLE and as such could be considered a syndrome in its own right, familial MTLE (FMTLE). They found that around one fifth of participants newly diagnosed with benign TLE were better described as having FMTLE. The reason this syndrome goes largely unrecognized, they argue, is because many of the symptoms (e.g. déjà vu) can be mostly perceived as normal. However, what this study highlighted which is possibly more relevant to establishing equivalence is the phenomenological differences (in terms of what people associated with “triggering” déjà vu) they defined between epileptic déjà vu, suspicious déjà vu experiences (akin to benign TLE but not significant to merit a diagnosis without further investigation) and normal “physiological” déjà vu. Normal experiences of déjà vu were more frequently associated with being in a certain situation (e.g. being in a new place) or with a certain modalities (e.g. initiating a certain movement) than suspicious and epileptic manifestations of déjà vu experiences. These (epileptic and suspicious) cases of déjà vu were characterized as prolonged (i.e. few seconds to several minutes), of a severe intensity occurring several times a year with possible occurrence in clusters. On the other hand, “physiological” cases of déjà vu were described as mild, fleeting and occurring yearly or less. (Though based on déjà vu
research in the wider population, we would expect “physiological” déjà vu to also occur more frequently than this.) Despite the proposed convergence on phenomenological differences between déjà vu in TLE and outwith this condition, this study noted poor inter-rater reliability when characterizing patients’ and relatives’ déjà vu experiences, which supports the point regarding our limited ability to diagnose these ambiguous non-lesional cases on the basis of their characteristics.

There are reasons to resolve this debate. First, experiencing déjà vu in this way or like these cases could be of diagnostic value. Second, it could help clarify one of the current conceptual issues within the field; that is whether déjà vu is driven by top-down processes (i.e. misfiring) or bottom-up processes (i.e. environment cues). The ability to distinguish healthy non-clinical déjà vu from those experiences that may be underpinned by different neurophysiological activity could be important for understanding the functional significance of déjà vu to the memory system, and to the resultant emotional and metacognitive reflections in people with epilepsy. However, as mentioned previously, our ability to establish differences between groups is significantly dependent on the measures used in the study design and currently, with respect to TLE, neuroimaging protocols are limited in their ability to provide differential diagnoses and self-report measures result in little correspondence between raters.

**Experimental Approaches in Broad Samples: Experiential Assessment.** All the above issues discussed in relation to natural déjà vu experiences reported at the time of the event or retrospectively should be considered in experimental work as well. When we ask participants in the lab to report experiencing déjà vu or make other memory judgments during analogues, similar study design considerations (e.g. question framing and retrieval query format) have been shown to impact participant responses or behaviour (Mill & O’Connor, 2014; Raposo, Frade, & Alves, 2016). With respect to reporting déjà vu in the lab and unreliable self-reports, Jersakova, Moulin, & O’Connor (2016) solicited déjà vu reports for a recognition memory task that was not expected to result in déjà vu. They used a continuous recognition task for which there was no theoretical basis for the elicitation of déjà vu (i.e. with no generation of the experience antecedents of false familiarity or novelty awareness), and which has never been associated with the experience. The expectation was therefore that the likelihood of déjà vu during the task should not ordinarily be elevated beyond the day-to-day likelihood of the experience—effectively nil in such a short experiment. Nonetheless, they found that repeated questioning about the experience during the task led to participants reporting déjà vu at high rates (unlikely attributable to a lowering of the decision-making threshold at which déjà vu would be reported given the total absence
of experimentally-generated déjà vu antecedents in the first place). Changing the questioning method from repeatedly asking participants whether they experienced déjà vu to one that required a free report when they experienced déjà vu reduced the overall number of déjà vu reports. Obtaining accurate reports (whether prospective or retrospective) hinges on adequately portraying the task demands, reducing the interpretative load on the participant and avoids constructing naïve beliefs in participants as much as possible (Conrad, Schober, & Coiner, 2007). This requires a balancing act of sorts. If we reduce the amount of subjectivity in the study by providing definitions of déjà vu and clear task demands, the assumption would be that we risk biasing participants into responding according to our hypothesis. The limited evidence we have on this point within déjà vu analogues would suggest that with regards to reporting rates, the definition makes no difference, but the task demands (i.e. question format) do prompt acquiescence bias in participants (Jersakova et al., 2016). This is backed up in retrospective reporting work as well, when we ask participants more regularly or offer a sense of rate information in the question and response format (e.g. regularly, weekly), then we reduce underreporting rates by encouraging participants to respond in a way that they perceive as social desirable rather than recalling their own experiences (Jersakova et al., 2016; Lugtig et al., 2016).

**Overcoming Methodological Challenges with a Convergent Approach**

The experiences we can be most certain represent déjà vu in the general population, are those that naturally occur within this population (rather than those analogues we experimentally generate in them, or indeed those of patients). Whilst the source experience is the one most déjà vu researchers wish to measure directly, the effect the intervening time between the experience and its report, is, as yet, unknown. Based on the literature discussed above though, we can assume that inaccuracy, in the form of reconstructive bias, or forgetting, will distort retrospective account. Even if these distortions were to be fully elucidated, such is their likely heterogeneity that it is impossible these effects could be accurately removed (to the right degree, from the right people’s reports). This leaves a residual issue that also affects some prospective research, as we outline below.

Given the importance of naturalistic déjà vu experiences, there has been some attempt to use them as a gold-standard against which other experiences are measured. Indeed, within our lab, this is something that has been done to establish the correspondence between experimentally-generated analogues and ‘true’ déjà vu both in experimental work (e.g. Urquhart et al., 2014; Jersakova et al., 2016; Cleary, Huebert, McNeely-White, & Spahr, 2019; Sugimori & Kusumi, 2014) and in the advice we have dispensed more broadly
(O’Connor & Moulin, 2010). For example, Urquhart et al. (2018) used these comparisons to describe accounts of their experimental analogue as more specific to the eliciting stimuli, shorter in duration, and less intense when compared account of “a “typical” déjà vu experience” (page 4 of online manuscript). Whilst we argue that this comparison provides a valuable additional source of information compared to taking an isolated account of the analogue at face value, the problem here is that the gold-standard is the same flawed retrospective measure we have critiqued above. Indeed, by minimising the difference between the experience we generate in the lab and these retrospective accounts, we may actually steer our analogues towards generating a caricature of the experience that most closely resembles the folk understanding of what déjà vu ‘should’ be.

Reporting naturalistic déjà vu experiences as they occur, prospectively, would reduce concerns around memory distortion and biasing reports by making use of responses based on experiential knowledge and minimally reliant on reconstructed memory (Conner & Barrett, 2012). With this in mind, we look to a small body of earlier work that has used prospective reporting on special samples—special because they were motivated to keep a diary of their experiences or because they were special samples that warranted such close clinical attention that their accounts of real-time déjà vu-like experiences could be transcribed as they were given. Heymans (1904, 1906; cited in Sno & Draaisma, 1993) asked students to record their experiences of déjà vu and depersonalisation in two studies, whilst Leeds (1944) kept personal diaries of his own frequent déjà vu experiences (one every 2.5 days). It is difficult to draw meaningful conclusions from these studies due to the lack of standardised information recorded alongside the accounts. For example, though Leeds (1944) provided exceptionally detailed accounts of 144 experiences over a 12 month period (with ratings of intensity, duration, his physical surroundings, antecedent physical states and behaviours for each experience) there has been no comparable record from any other individual or sample. A sample in which more standardised physiological information was recorded was the patients experiencing ‘dreamy state’ reported by Penfield & Perot (1963). Not only did the authors transcribe and analyse the patients’ subjective experiences via verbal descriptions of their memory sensations as they experienced them, they also kept records of the stimulation points that elicited the experiences. That these physiological stimulations could be revisited, to test whether subjective experiences could be re-evoked and therefore localised to a physiological source, highlights the power of convergent measurement of a range of measures alongside the subjective experience at the centre of interest. Whilst Penfield and Perot’s patients were special samples undergoing brain surgery, and experiencing something akin to déjà vu, we believe there is potential for today’s déjà vu
researchers to draw inspiration from their methods, gathering prospective subjective, physiological and situational data to build the most accurate picture of naturalistic déjà vu.

Three characteristics set apart the approach we propose from those naturalistic approaches of Heymans (1904, 1906; cited in Sno & Draaisma, 1993) and Leeds (1944), and the clinically relevant work of Penfield and Perot (1963):

1. **Standardisation.** The déjà vu research community, galvanised by Brown’s (2003, 2004) reviews of the early literature, has in the past two decades begun to communicate, collaborate and publish together in ways that lend themselves to large-scale collaboration. Much of the literature reviewed by Brown had to be examined in a piecemeal manner; the items used to assess even basic facets of the experience were unique to each piece of research, making comparison difficult. We can overcome this by taking a unified approach such that research groups across the world negotiate joint priorities, and set in place a framework (e.g. probed experiences, demographics and situations, item wording) for the standardised prospective measurement of déjà vu.

2. **Large-scale, long-term.** The infrequent occurrence of déjà vu means that a sample of 100 adults drawn from the general population and brought into the laboratory for an hour each, will likely yield no spontaneously occurring naturalistic déjà vu reports. Any attempt to prospectively record déjà vu experiences would require a large-scale long-term approach, such that those involved would not be participating in an experiment, but going about their daily life and recording the events of interest as they happen. This citizen science approach has been successful in a range of domains that require long-term participation (e.g. air quality monitoring by 20,000 members of the public in Flanders; reported in Irwin, 2018) or the reporting of relatively unusual events (e.g. magpie sightings in South Australia; Zeegers et al., 2012) and has been facilitated in recent years by the willingness of the general population to participate using their mobile phones.

3. **Ubiquitous Technology.** Mobile apps lend themselves to simple, semi-automated data gathering. Events can be recorded on simple interfaces, and complemented with other automated data: a range of sensors gather highly localised information (e.g. time, location, audiovisual recording), whilst connectivity with other apps provides access to situational information (e.g. weather). Perhaps most excitingly, mobile phone platforms such as Apple and Android are providing ways with which physiological data from wearable devices (e.g. smart-watches generating electrocardiograms from optical sensors and smart-glasses with the potential for
simple electroencephalogram from a small number of channels) could in the future be incorporated into event records.

The benefit of a large-scale, long-term, standardised approach is that there would be no need to focus on specific groups (e.g. trading off the frequency of the experience against the ‘representativeness’ of the sample in which it could be recorded). Anyone who wished to contribute could participate, and their specific demographic characteristics and behaviours (e.g. medical conditions, substance use etc.) could be used to build pictures of the experience in different demographic groups. Not only would this allow the research community to understand the links between demographic, situational and physiological variables on subjective and objective characteristics of the déjà vu experience, it would in doing so, allow us to establish whether or not the differences we assume from the individual studies of the general population and special samples are evident in these prospective data. Separate research groups with differing priorities could continue to focus on recruiting their own special samples, but with the advantage of having access to all the data, from all participants, special or otherwise, gathered from the collaborative project.

The big picture here is that the study of déjà vu has reached a point where a collaborative approach stands to benefit the field. Within this collaboration, researchers would continue to investigate the experience with the priorities we have grown interested in, but with the majority of the problems of measurement addressed by methodological change, or in-built control. Research on broad samples would minimise bias by transitioning from retrospective to prospective reports. Baseline data would be useful not only for those interested in clinical samples, but those working on experimental analogues. The situation described above, in which prospective déjà vu analogue reports are compared to retrospective naturalistic déjà vu accounts would be replaced by a comparison of like with like, with the same standardised battery of items used across all situations.

How exactly could this be achieved? Large-scale collaborative projects, particularly those aimed at dealing with issues of reliability in the psychology literature, have already been successful (e.g. “Many Labs” projects; Klein et al., 2014; 2018). Within these projects, many labs from across the world have planned and conducted replication studies, choosing first which effects to investigate—prioritising and standardising—before recruiting for and conducting a number of studies. This collaboration has been catalysed by the identification of a problem within the field, and enabled by real-time internet tools and uniform deployment of experiments designed to agreed-upon standards. The issue of how this could be adapted and rolled out to citizen scientists is more challenging. A fully-fledged déjà vu app would be
the most coherent way to implement this, though it would likely suffer sampling biases and could struggle to reach engagement that could be called large-scale. Integration with other mobile ‘health’ suites could vastly improve reach, but would be unlikely given the niche even with psychological research that déjà vu occupies, and would also likely compromise the researchers’ control of the items and data themselves. This is undoubtedly a big hurdle, requiring collaboration beyond the sphere of psychology research, but which if solved in the same planning stage at which the issues of scientific priorities and standardisation were agreed, would make this an exceedingly powerful approach for the study of déjà vu.

Conclusions
The current chapter has reviewed the approaches déjà vu research has taken from its early references and exploration in literature to now, where the study of the phenomenon spans various perspectives, methods and disciplines. The diversity and divergence of approaches for studying déjà vu may, in part, be a consequence of the various populations associated with experiencing déjà vu and the lack of behavioural marker or consequence to verify the occurrence of déjà vu in the general population. We set out the three emergent research priorities for the study of déjà vu and evaluated the methods that have been used to investigate the déjà vu experiences of their relevant samples. With consideration of exploring naturalistic experiences in broad samples and in the absence of behavioural verification of déjà vu, the study of the experience in this approach has relied on the retrospective report. From the review of similar literature in health and clinical research, the use of these methods is highly unreliable and susceptible to bias. Likewise, approaches investigating observable experiences in special samples share some of the self-report methods used in broad samples and have tended to focus on phenomenological differences in déjà vu as part of temporal lobe epilepsy. Recently, in this approach, there has been a drive to categorise déjà vu experiences on the basis of clinical relevance that they might have rather than what the phenomenological spectrum of those experiences might be able to offer our understanding of déjà vu in the general population. Most recently, observable experiences in broad samples have been generated through the development of lab analogues and demonstrated our ability to bias déjà vu reports within the context of experimental studies. One of the main shortcomings of having these disparate approaches each with their own priorities is that findings such as what factors influence déjà vu, in what sample and in what context rarely extends across those disciplinary or theoretical divides.

Finally, we argue that the early neuropsychological approaches of Penfield assessing the experience of déjà vu as it came to fruition in the individual is an integrative approach that has been largely overlooked in contemporary research. In déjà vu, we have a quirky memory
experience that is phenomenologically rich but demonstrably prone to bias regardless of the research methods used to probe the experience. The use of real-time or prospective reporting of déjà vu experiences through online and mobile technologies in broad and special samples alike would help overcome many of the biases found in the various approaches we have discussed. It would also help bridge previous divides between cognitive and clinical approaches by offering averaged baselines of prospective déjà vu experiences allowing for better comparison between general and special samples. Converging on the study of déjà vu in this way, while it would require a large scale and long-term approach, would mean the investigation of the experience would be less restrained by the biases found across research approaches and prioritise the subjectivity of the experience in our understanding of naturalistic déjà vu.
References


Augustine. (400AD). Book XII, Chapter XV. In *On The Trinity*.


