Original Papers

Daisuke Nakamura*

Methodological considerations in studying awareness during learning: Part 1: Implicit learning

Abstract *Methodological problems of how awareness during learning should be measured have been extensively discussed and investigated in cognitive psychology. This review considers; 1)whether amnesics can perform implicit learning tasks at a similar level to normal controls, 2) whether differences in instructional orientations create dissociations in performance in tests of implicit and explicit knowledge, and 3) whether participants can retrospectively verbalise the learning outcomes. The paper concludes that; (1) amnesics' implicit learning abilities differ from the normal controls, (2) instructions on implicit learning do not guarantee the occurrence of implicit learning, and (3) objective and subjective awareness measures used in the literature face inherent problems and so the awareness controversy remains unsettled.*

Key words: *awareness, implicit learning, review, methodological problems*

How can we measure awareness during learning in a sufficient and reliable way? This question remains to be answered despite extensive empirical efforts in the implicit learning literature (see Perruchet, 2008 and Shanks & St. John, 1994, for reviews). The aim of the paper is therefore to consider the merits and demerits of these empirical methods and thus to provide the state of the art in terms of measuring awareness and suggest possible future directions for research. In doing so, the paper first considers whether the results from amnesics' performance on implicit learning are generalisable to the normal population and whether participants follow instructions on implicit learning. Finally, the advantages and disadvantages of each objective and subjective measure of awareness are compared since multiple employment of these measures (e.g., the post-decision wagering) has become a recent trend in the literature.

Some research has tried to separate implicit learning from explicit learning in terms of employing different populations such as amnesics who are argued to have impaired recognition memory or in terms of different learning instructions on implicit/explicit learning. However, neither approaches could show learning without awareness: amnesics show different performance from normal participants even on implicit learning (Channon et al., 2002; Curran,

1997) and thus results from the former population cannot be generalisable to the latter population. Differences in learning instructions also cannot guarantee that participants simply follow the instructions and results confirmed this is the case.

Therefore, it is important to measure participants' awareness during learning with some indexes of awareness. The verbal report, whether implemented concurrently or retrospectively, has been criticised since it cannot rule out the possibility that participants develop unverbalisable yet conscious *phenomenal* awareness (Block, 2005, 2007). The verbal report can capture only higher-order *access* awareness that participants can access to contents of at best. The objective measures such as grammaticality judgment (GJ) recognition, or strategic control in the process-dissociation paradigm, on the other hand, may measure phenomenal awareness because of its forced-choice nature: participants are not required to verbalise the contents of their minds. Objective measures, however, particularly if they are employed as dichotomous measures, are able to capture only coarse-grained states of awareness (aware/unaware).

Subjective measures such as confidence rating, on the other hand, can capture subtle or graded states of awareness since they are often implemented with continuous scales or

* Affiliation: Aoyama Gakuin University, 4-4-25, Shibuya, Shibuya-ku, Tokyo. JAPAN150-8366 Mail: hamuhamuhamihami@ yahoo.co.jp

with variable response options. Therefore, they can capture more graded states of awareness during learning and have thus become a recent trend in the implicit learning literature, though their subjective nature has inherent problems such as reluctance to claim confidence.

Before presenting detailed discussions of the above issues, we begin with a brief conceptual introduction to implicit learning.

Implicit Learning

There have been a wide variety of definitions of implicit learning and thus the term is polysemous (Frensch, 1998). One of the original definitions was provided by Reber (1993, p.12):

A situation-neutral induction process whereby complex information about any stimulus environment may be acquired largely independently of the subjects' awareness of either the process of acquisition or the knowledge base ultimately acquired.

This definition claims that both (1) the learning processes and (2) the resultant knowledge are implicit, the special concession "largely independently" implying that both the implicit learning processes and implicit knowledge can interact with explicit learning processes and explicit knowledge. Reber can be said to hold the weak interface model though the literature does not interpret his view in this way, nor has he specified the conditions in which interaction happens.

Implicit learning has been studied in various paradigms including artificial grammar (AG) learning (Reber, 1989a, 1993), sequence learning (or serial reaction time [SRT] tasks, e.g. Nissen & Bullemer, 1987; Richard, Clegg, & Segar, 2009), and dynamic system control (Berry & Broadbent, 1984). In the typical sequence learning paradigm (Nissen & Bullumer, 1987), a light appears in one of four locations on a screen and is presented sequentially up to 10 times, one by one. Then, this fixed 10-light sequence is repeated many times. Participants are simply required to press a corresponding key, which is located below each of the four locations. Compared with a control group that sees the light generated in a random order, the reaction time (RT) of the experimental group is significantly reduced as the experiment goes on.

In the dynamic system control paradigm, which has been less frequently adopted than the other two paradigms, participants are required to control complex computer-implemented systems. For instance, Berry and Broadbent (1984) used a sugar production task where participants, supposedly in charge of a sugar production factory, were required to control the production rate of the sugar by changing the size of the work force (i.e. the number of workers). Each target production rate was generated by a fixed equation,

but only the current work force and production rate were provided to participants.

The most widely used paradigm is artificial grammar learning (AGL) wherein digit sequences (e.g. XXVX) are generated by a finite state Markov grammar. During training, subjects are told simply to memorise digit sequences. In the test, after being told about the existence of the rule underlying the digit sequences, they judge the grammaticality of the new digit sequence that is produced by the same rule as the training digit sequences.

Typically, participants' awareness is measured by a verbal report or a written questionnaire. Participants cannot verbalise the rules underling the digit/number sequences (AGL and SRT task) or the relation between the production rate and the size of the work force (dynamic system control) but nevertheless can indicate correct grammaticality judgment at an above-chance level. Therefore, the acquired complex knowledge is argued to result from implicit learning and to be tacit, i.e. unavailable to conscious awareness. For instance, Reber (1976) compared the explicit learning condition, where the explicit instructions to memorise items as well as to search for underlying rules were given in the training session, and the implicit learning condition ,where the instruction simply memorise to items was imposed. He found that participants under the implicit learning condition outperformed participants in the explicit learning condition with regard to their accuracy in remembering training items. Moreover, learners under the implicit learning condition significantly outperformed learners under the explicit learning condition on grammaticality judgment of the new transfer items.

However, there has been extensive criticism of the nature of implicitness during learning and/or acquired knowledge (see Perruchet, 2008 and Pothos, 2007, for reviews). The criticism seemingly stems from the methodological flaws in the experimental design (e.g. whether participants in the implicit learning condition simply memorise the digit sequences during training or actively search for the underlying rule). The paper presents an overview of these methodological issues in the next section.

Methodological Considerations of Awareness during Learning

No consensus has been established with regard to the possibility of learning without awareness despite extensive research on implicit learning. This is partially owed to the confusion between whether unconsciousness refers only to the learning processes (i.e. whether participants simply try to memorise the digit sequence without consciously searching for underlying patterns in the case of AGL) or both the learning processes and the acquired knowledge (i.e. whether participants simply try to memorise the digit sequence without consciously searching for the underlying rule **and** are aware of acquired knowledge of the rule).As

Frensch (1998) observed, if the former is the case, the implicitness of the learning processes itself is measured by testing the implicitness of the acquired knowledge in the implicit learning literature. The implicitness of the learning processes could be only 'inferred' from the state of the learned knowledge in this case, however: the possibility that knowledge becomes implicit for other unknown reasons cannot be eliminated.

On the other hand, when the implicitness refers to both processes and results as can be seen in the definition of implicit learning provided by Reber (1989a, 1993, quoted above), there is no 'independent' test for the implicitness of the learning processes, though implicit/explicit learning processes are often operationalised as the provision of different tasks for implicit and explicit learning (e.g., Reber, Walkenfield, & Hernstadt, 1991) or different instructions for the same task (e.g., Gebauer & Mackintosh, 2007). There is no guarantee, however, that participants will follow instructions (see below for detailed discussion). Nevertheless, evidence for learning without awareness seems to be primarily drawn from the findings that:

(1) Amnesic patients can perform implicit learning tasks at a similar level to normal controls.
(2) Differences in instructional orientations (e.g. intentional rule-search vs. incidental memorisation) create dissociations in performance in tests of implicit and explicit knowledge.
(3) Participants cannot retrospectively verbalise the learning outcomes.

Subsequent sections deal with these issues in detail.

Can Amnesics Perform Implicit Learning Tasks at a Similar Level to Normal Controls?

Amnesics have often been described as those who have severe impairment of their explicit declarative memory but whose implicit non-declarative memory remains relatively well-preserved. If they cannot recognise or recall the underlying grammar but nevertheless show above-chance performance in AGL or SRT tasks, it has been argued that they show implicit learning. Knowlton and Squire (1994, 1996) found just that. They compared both normal participants and amnesic patients as regards AGL. They found that both groups could judge the grammaticality of the new transfer sequences. Since amnesics' recognition of fragmentary knowledge (or chunks, e.g. XVX) was marginally significant ($p = 0.06$, experiment 2 in Knowlton & Squire, 1996), they concluded that amnesic patients could not rely on the impaired recognition memory for the chunks.

Subsequent experiments on amnesics' implicit learning provided counter-evidence for Knowlton and Squire's interpretation, however: (1) amnesics had immediate recognition memory similarly to normal participants; (2) the same amnesics were repeatedly recruited in a series of Knowlton and Squire's experiment, which might create above-chance performance on AGL; (3) when the stimulus domain became complex, amnesics showed impaired implicit learning abilities. These are described below.

First, 'impairment' does not indicate total 'lack' of recognition memory and explicit learning processes. Sometimes, 'Amnesia is not a failure to notice a novel conjunction; it is a failure to consolidate an explicit memory as a result of noticing' (Ellis, 2005, p.319). There is some evidence for this. Della Sala, Cowan, Beschin, and Perini (2005) found that anterograde amnesics could recall the contents of stories just as normal controls did, as reflected in the equal performance on the immediate recall task. Their delayed recall was significantly lower than the controls', however, suggesting that amnesics have explicit memory and can encode information into long-term memory. Where they fail, as Ellis (2005) observes, is in consolidation of and subsequent retrieval of the encoded information.

Second, it should be noted that in a series of Knowlton and Squire's experiments (e.g., Knowlton, Ramus, & Squire, 1992; Knowlton & Squire, 1994, 1996), *the same* amnesics were recruited repeatedly, possibly because of the low availability of amnesic patients, though the intervals between experiments were longer (e.g. five months). Therefore, the possibility that they had been repeatedly trained to learn AG and this created the apparent learning effects cannot be excluded.

Third, there is also evidence that amnesics, when a structural domain to be learned becomes complex (e.g. biconditional grammar), show reduced implicit learning abilities compared with their normal counterparts (Channon et al., 2002; Curran, 1997). Curran (1997) compared anterograde amnesics with corresponding normal controls on SRT tasks. Two types of sequences were used: (1) the first-order predictive (FOP) sequence (e.g. A-B-A-D-B-C-D-C-A-D-B-C) whereby each element (e.g. A) could be followed by one element (e.g. D) 67 % of the time or to a lesser extent by another element (e.g. B) but by not the other elements; and (2) the second-order predictive (SOP) sequence (e.g. A-B-A-D-B-C-D-A-C-B-D-C) whereby each element can be followed equally by all the other elements. Frequency information, that is, the transitional probability of pairwise elements (e.g. possibility of D given A) is only effective in the former sequence. The results showed that both groups learned both types of sequences as reflected in decreases in RTs in the later trials. Magnitude of learning was greater in the SOP than in the FOP sequences for the normal controls but this was not the case for the amnesics. That is, only the normal controls could use predictive information on the test stimuli. These results indicate that amnesics do not have similar implicit learning abilities to their normal peers.

The lack of abstract or higher-order knowledge in amnesics' implicit learning was also confirmed by Channon

et al. (2002), using AG. The biconditional grammar used in their studies imposed constraints on two non-adjacent positions (e.g. first and fifth positions), such that if one element contained D, then another should be F, and the same was true for G/L and K/X. For example, DFGXFDLK was grammatical whereas LFGKKDLX was ungrammatical. The authors compared amnesics' learning of this biconditional grammar with the normal controls' learning by means of orthogonal manipulations of the grammaticality and the associative chunk strength (e.g. a percentage of a test item containing old bi- or trigrams presented during training). Their GJs, with a six-point confidence scale and immediate recall of the six letters, were tested. In both groups, the grammaticality had no effects but the chunk strength did, and in such a way that both groups judged more items with higher chunk strength as grammatical than items with lower chunk strength. The results of the confidence rating revealed that there were no differences between correct and incorrect judgment, suggesting that learning is primarily implicit. In addition, amnesics' recall of the six letters used in the experiments was significantly worse than that of the controls, suggesting that amnesics' explicit memory is reduced, as the authors argued. Therefore, Channon et al.'s (2002) studies show that what amnesics learn during AGL is chunk information not abstract knowledge of the underlying rules.

These results, taken together, suggest that implicit learning is not very robust in the face of various neurophysiologic damage, as often assumed in the literature (Reber, 1989a, 1993), and amnesics in particular are not good at learning complex stimulus domains such as bidirectional grammar. Amnesics' implicit learning abilities are significantly reduced such that they can learn only knowledge about co-occurrence of stimuli (chunks). In this sense, the amnesics differ from the norm even in their implicit learning abilities, let alone explicit learning abilities. Therefore, the results of the amnesics' AGL cannot be taken as positive evidence for learning without awareness in the normal population. The next sections review experiments recruiting only the normal population.

Do Differences in Instructional Orientations Create Dissociations in Performance in Tests of Implicit and Explicit Knowledge?

In the typical AGL experiment (e.g., Gebauer & Mackintosh, 2007; Reber, 1967), two different instructions are given to two groups of participants as a between-participants variable: (1) intentional: to search for the rules underlying training stimuli, and (2) incidental: to memorise or simply see the training stimuli. Alternatively, only (2) is given with additional explicit learning tasks such as paired-associates learning (Brown, Aczel, Jimenez, Kaufman, & Grant, 2010), a forced-choice, series solution problem task where participants complete an alphabetical sequence (e.g. ABCBCDCDE, D or C; Reber, et al., 1991) as a within-participants variable. In the former case, the GJ task with some measures for awareness are provided as tests for the acquired knowledge and the scores of the explicit learning group are compared with those of the implicit or incidental learning group. On the other hand, in the latter repeated-measure design, comparison of the same participants' performance across implicit and explicit learning tasks itself serves as the test for awareness (an additional retrospective verbal report is sometimes provided, e. g. Robinson, 2002, 2005).

Both between-participants and within-participants designs for implicit learning have several inherent as well as general flaws, however: namely, (1) task differences instead of differences in learning processes (Gebauer & Mackintosh, 2007); (2) task contamination (Segar, Prabhakaran, Poldrack, & Gabrieli, 2000); and (3) ineffectiveness of instructions. These are considered below.

First, some studies employed different tasks for implicit and explicit learning. For example, Reber et al. (1991) operationalised explicit learning as a forced-choice problem-solving task (as described above) and implicit learning as memorisation of the digit sequences as in typical AGL within a repeated-measure design. The results showed the independence of implicit learning from intelligence measured by WAIS-R (short form) but significant positive correlations between explicit learning and intelligence. Gebauer and Mackintosh (2007), however, pointed to the possibility that the results might come from differences in tasks, not from differential susceptibility of implicit and explicit learning to individual differences in intelligence.

Second, the repeated-measure design also has an inherent problem: implicit learning and explicit learning tasks may contaminate each other.[1] Obviously, presenting explicit learning first and then implicit learning is problematic since participants may bring some explicit learning strategies (e.g. conscious rule search) driven by the first explicit learning task into the subsequent implicit learning task. The reverse order might suffer from the same problem, however: participants continue to try to memorise stimuli in the explicit learning task, which hinders or facilitates explicit learning (see Perruchet & Pacteau, 1990, for positive effects of memorised chunks). Segar et al. (2000), to justify their between-participants design for AGL, commented on their pilot study that had employed the repeated-measure design: '…pilot testing indicated that participants often adopted the same decision strategy for both recognition and grammatical judgment, even when they were instructed to perform the tasks differently' (p.285).

---

[1] However, the within-participants design has a positive aspect in a replication study. (see Robinson, 2009).

Third, some studies employed the between-participants design by providing different learning instructions for the same task in order to avoid the issue of task differences. For instance, Gebauer and Mackintosh (2007), in their comprehensive studies, instructed participants in the explicit learning condition to search for general as well as micro rules that constrained repetitions or bigrams (typical two-digit pairs such as XV) or trigrams (typical three-digit sequences such as XVX). Participants in the implicit learning group, on the other hand, were instructed to memorise the digit sequences generated by the same grammar as in the explicit learning condition. Even in this between-participants design, however, it seems possible to assume that implicit learning *instruction* cannot guarantee that participants followed the instruction by simply memorising the training digit sequence. Although there seems to be no direct evidence for this in the implicit learning literature, there is evidence in the implicit natural language learning literature. Robinson (1997) provided this. In his experiment, Japanese second-language learners in the implicit learning condition were instructed to memorise training sentences of various English constructions. The post-experimental awareness questionnaire revealed, however, that of these (N = 24), nine claimed to look for underlying rules and six could actually verbalise the rules (see Nakamura, accepted, for detailed methodological considerations of awareness in the second-language acquisition literature).

How can we solve the above methodological problems? If different learning instructions for the same AG are employed, different levels of awareness should be measured by a questionnaire and plotted against each learning condition and frequency distributions of each cell should be compared using X2 or removal of those who do not follow the instructions. Unfortunately, the scores on the awareness questionnaires are averaged within the particular group in the typical study and those who do not follow the instruction cannot be removed from the analyses. Therefore, individual differences in the development of awareness should be assessed in order to confirm that participants in the implicit learning condition do not consciously test hypotheses on the training stimuli. Otherwise, differences in instructional orientations cannot guarantee that participants actually engage in an expected learning mode.

A more promising way to operationalise implicit learning is perhaps to employ the *single* learning condition with implicit learning instruction and then use some measures of awareness classify participants into those who engage in implicit learning and those who engage in explicit learning, as the preceding paragraph suggests. This single-task design can avoid the problems of task contamination, task differences, and ineffectiveness of instructions discussed above. An important problem persists, however: how can we measure awareness during learning sufficiently? This topic is dealt with in the next section.

Can Participants Retrospectively Verbalise Learning Outcomes?

In the typical AGL and SRT tasks, awareness is measured by the retrospective verbal report (Reber, 1967). Since then, the insufficiency of the verbal report has been extensively criticised (Brody, 1989; Eriksen, 1960; Holender, 1986; Shanks, 2005; Shanks & St. John, 1994, though see Runger & Frensch, 2010 and Weiskranz, 1997, for different claims). This section points out the insufficiencies of the verbal report with particular reference to Shanks and St. John (1994) and then introduces alternative measures of awareness, which are reviewed in subsequent sections.

Shanks and St. John (1994) identified two criteria that must be met by studies claiming to demonstrate learning without awareness: (1) the information criterion, i.e., whether awareness assessed by tests is indeed responsible for task performance, and (2) the sensitivity criterion, i.e., whether tests of awareness are indeed sensitive to all necessary relevant explicit knowledge. The verbal report is typically assessed at the end of the experiment, that is, after the test session. This means that the awareness of time 1 (learning) is inferred from the results of the verbal report at time 2. Obviously, participants' contents of awareness decrease or even fade away at time 2 (Shanks & St John, 1994). In other words, the retrospective verbal report cannot meet the sensitivity criterion.

To make the verbal report more sensitive to the contents of awareness during learning, researchers developed concurrent verbal reports such as the think-aloud protocol (Ericsson & Simon, 1993). A concurrent verbal report, that is, when someone verbalises the contents of the acquired knowledge or the process of learning while engaging in the learning task, cannot avoid the problem either, because it can interfere with the learning task all too easily (Jourdenais, 2001; Shanks & St. John, 1994).

Moreover, the verbal report, whether it is implemented retrospectively or concurrently, cannot exhaustively measure the contents of conscious minds. Ned Block (2005, 2007) drew a distinction between *phenomenal* and *access* consciousness. Phenomenal consciousness refers to awareness of but not access to contents of mental states whereas access consciousness is awareness of and access to contents of mental states (cf. reflective consciousness, which is verbalisable access consciousness). Phenomenal consciousness could remain without access or retrospective consciousness. This raises the possibility that participants still have phenomenal but not access consciousness during implicit learning. The verbal report by definition cannot measure phenomenal consciousness though it might be able to measure access consciousness.

Therefore, the verbal report cannot rule out the possibility that participants are phenomenally aware of the learning processes and/or outcomes either at time 1 (learning)

or time 2. On the other hand, explicit knowledge measured by the verbal report at time 2 might be evidence for explicit learning. This is because participants still hold at least explicit knowledge measured by the retrospective verbal report without forgetting it.

In order to tackle the methodological problem of measuring awareness, a number of measurements have been developed (Seth, Dienes, Cleermans, Overgaard, & Pessoa, 2008). These include objective measures such as GJ, recognition, generation, and strategic control such as Jacoby's (1991) process-dissociation procedure (e.g., Destrebecqz & Cleermans, 2001; Wilkinson & Shanks 2004), and subjective measures such as the mere exposure effects (Manza, Zizak, & Reber, 1998), confidence rating (e.g., Dulany, Carlson, & Dewey, 1984; Perruchet & Pacteau, 1990), and wagering (e.g., Dienes & Seth, 2010). These are administered either dichotomously or continuously with some scales. The following sections consider these.

Dichotomous GJ, recognition, and generation

An alternative way to measure awareness more sensitively is by forced-choice tasks that employ some dichotomous response. The objective measures of awareness are GJ, recognition, and generation. These are dichotomous yes-no type tests when they are administered without continuous scales. Recognition tasks, which have been extensively used in the memory literature, require participants to judge whether a stimulus is one they saw during training or not. Generation tasks are often used in the SRT task (Destrebecqz & Cleermans, 2001; Norman, Price, Duff, & Mentzoni, 2007; Shanks, Rowland, & Ranger, 2005; Wilkinson & Shanks, 2004). One of the generation tasks used in these studies is to predict the next position given five stimuli. For example, Norman et al. (2007) asked participants to indicate whether the sequence produced was 'less confident' or 'more confident' in a trial-by-trial way, which was approximate to the concurrent testing of awareness.

GJ, recognition, and generation are more objective than the verbal report because participants are simply required to categorise stimuli into two categories (e.g. *seen* and *unseen*). In addition, the objective measure is more sensitive than the verbal report since participants can respond on the basis of phenomenal awareness whereas the verbal report cannot capture phenomenal awareness.

These measures also have several pitfalls, however. First, GJ has been used as the test for implicit knowledge (e.g., Reber, 1967). Reasoning in the typical implicit learning reflects this: participants cannot verbalise the rules but nevertheless can judge the grammaticality of the test stimuli and therefore the resultant knowledge is implicit. GJ, however, is the test for explicit knowledge, particularly when it is administered without time constraints (R. Ellis, 2005; R. Ellis Loewen, Elder, Erlam, Philip, & Reinders, 2009). To judge the grammaticality of the test stimuli, participants

must engage in hypothesis testing with mental effort. The psychometric studies confirmed the following: scores on the untimed grammaticality judgment were loaded on the explicit knowledge factor (R. Ellis, 2005, 2009; Ellis & Loewen, 2007). Therefore, the fact that participants can judge the grammaticality at an above-chance level itself indicates the possibility that they engage in the explicit learning processes, which results in explicit though unverbalisable or phenomenal knowledge of the underlying grammar. Poznanski and Tzelgov (2010) also used GJ as a test for 'intentional' retrieval in the AGL task.

Second, the dichotomous discrimination without scales is too coarse-grained to capture the subtle nature of participants' awareness. For instance, participants might develop 'graded' (phenomenal) awareness but for some slight awareness of their mind is not enough for them to give 'Yes, I saw it during training' or 'Yes, I am more confident' responses. In other words, each participant's threshold of awareness is different, and without continuous scales, the dichotomous measure cannot deny the possibility that some participants have developed a lower level of awareness (see Wierzchon, Asanowicz, Paulewicz, & Cleeremans, in press, for a related argument).

Strategic control

The strategic control measures of awareness are based on the process-dissociation procedure (PDP),originally developed by Jacoby (1991, see also Yonelinas & Jacoby, 2012, for the latest review). According to Jacoby (1991), implicit (or automatic) and explicit (or controlled) processes are contaminated in any tasks (the process (im)purity problem; see also Frensch, 1998). In order to avoid contamination, two conditions are compared in the PDP: (1) the inclusion condition (I): participants are asked to *include* explicit knowledge learned during training, and (2) the exclusion condition (E): participants are asked to *exclude* such knowledge. The rationale behind the PDP is that when participants produce explicit knowledge, even following the instruction to exclude it in the exclusion condition, this probably reflects implicit knowledge, particularly when I equals E but is larger than some baseline performance (B). By qualitative dissociation of the implicit and the explicit processes within a particular task, the PDP aims to dissociate the effects of both processes.

This, or modified versions of the PDP, are utilised in the AGL and SRT tasks (Destrebecqz & Cleermans, 2001; Dienes, Altmann, Kwan, & Goode, 1995; Fu, Dienes, & Fu, 2010; Norman et al., 2007; Wilkinson & Shanks, 2004). In Destrebecqz and Cleermans's (2001) study, participants' explicit knowledge was measured in 96 trials of free generation of the learned sequence in both the inclusion and the exclusion conditions. Irrespective of the interval between each training sequence (0ms vs. 250ms), participants generated more old chunks that had appeared during training in

the inclusion condition. Conversely, when no intervals were available, they produced more old chunks in the exclusion condition despite the exclusion instruction than when a 250ms interval was available. Since subsequent analyses over data in the six-point recognition test mirror the results obtained by the PDP (participants could discriminate old from new chunks only in the 250ms interval condition), the authors concluded that learning obtained in the exclusion and 0ms interval condition was implicit in nature.

Wilkinson and Shanks (also Shanks, Wilkinson, & Channon, 2003), however, failed to replicate the results of Destrebecqz and Cleermans (2001). Wilkinson and Shanks (2004: ex 1) provided the inclusion and the exclusion conditions as a between-participants variable rather than a within-participants variable as in Destrebecqz and Cleermans (2001). This was because the inclusion-exclusion conditions as a repeated experimental design could confound the results with other variables such as change in attentional alertness. The results showed that the number of old chunks generated in the free-generation task was larger in the inclusion group than in the exclusion group (thus I > E). Moreover, there were no differences between old chunks and other incorrect chunks (baseline) in the exclusion condition whereas the number of the old chunks was significantly larger than that of the other incorrect chunks in the inclusion group (thus I > E = B). Norman and colleagues (Norman et al., 2007; Norman, Price, & Duff, 2006) also failed to replicate Destrebecqz and Cleermans (2001). In their experiment 2, Norman et al. (2007) extended the 250ms interval to 1,000ms in order to maximise differences in awareness between the short and long time intervals. Irrespective of differences in the interval, however, old chunks were generated more in the inclusion than in the exclusion conditions. Using the original 250ms interval, Norman et al. (2006) found the same results as in Norman et al. (2007). Therefore, learning without awareness is largely impossible in the PDP procedure.

One advantage of using the PDP is that the problem of the task differences can be avoided because the PDP in the implicit learning literature usually employs instructional differences for the same task instead of task differences. Another is that the problem of task contamination can be resolved by statistical calculation of inclusion and exclusion performance. The PDP has also been criticised for its particular methodological problems, however: (1) exclusion errors do not reflect effects of unconscious processes but result from external factors such as motivation; (2) exclusion instruction is too difficult to follow; and (3) when there are no exclusion errors, the magnitude of the unconscious processes cannot be estimated (Fisk & Haase, 2006; Snodgrass, 2002; Snodgrass & Shevrin, 2006; Visser & Merikle, 1999).
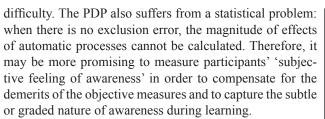
First, exclusion errors produced in the exclusion condition may be caused not because participants could not consciously exclude studied items but because they sim-

ply got bored, felt less confidence, or were less motivated. Remember that the exclusion errors come from the unconscious processes in the PDP: despite conscious effort to exclude, participants automatically include studied items. Visser and Merikle (1999) found that increases in motivation thanks to monetary incentive decreased exclusion errors in perceptual identification of a briefly presented word. In their Experiment 1, after masking, an English word was presented briefly (0, 50, or 250 ms). Then a stem with the first three letters was presented and participants were asked to complete the stem either by excluding or including the briefly presented word. In addition, the motivated group was told that if they made a mistake the participation fee would be reduced whereas the control group was not so informed. Visser and Merikle (1999) found that elevated motivation in terms of monetary incentive significantly reduced the exclusion errors. The results suggest that participants in the PDP were simply demotivated and thus made a mistake (though monetary incentive itself may cause an inherent problem, risk aversion. See 'Post-decision wagering' below). Snodgrass (2002) also raised the possibility that participants do not exclude the studied items unless they have higher confidence, whereas this is not the case for the inclusion task because 'inclusion instructions simply ask participants to provide their best candidate identification, regardless of their confidence in this response' (p.559). In other words, exclusion errors are simply apparent failure caused by external factors not by implicit processes.

Second, it is often the case that some participants have to be excluded from the data simply because they cannot follow the exclusion instruction (Debner & Jacoby, 1994; Hutchinson, Neely, Neill, & Walker, 2004, cited in Fisk & Haase, 2006). Fisk and Haase (2006) interpreted this as meaning that the exclusion task is too difficult compared with the inclusion task, and thus the exclusion task was confounded by 'task difficulty' (p.4250).

Contrary to the above case, there are participants who show no exclusion errors (Fisk & Haase, 2006). This leads to the third methodological problem: there is no way to estimate unconscious processes in the PDP. According to the PDP, conscious processes were calculated by subtracting exclusion from inclusion performance and thus conscious processes = inclusion – 0. Performance on the inclusion task, in the PDP, is a confound of automatic and controlled processes by definition. Therefore, the result is 'uninterpretable' (Fisk & Haase, 2006, p.4250).

In summary, dichotomous objective measures such as GJ, recognition, and generation cannot capture the subtle or graded nature of awareness during learning. Instead, these measures capture only a coarse-grained picture of the conscious minds at best. Strategic control, also dichotomous, was developed from the PDP and has inherent methodological problems: participants produce exclusion errors not because of automatic processes but simply because of external factors such as motivation, confidence or task

difficulty. The PDP also suffers from a statistical problem: when there is no exclusion error, the magnitude of effects of automatic processes cannot be calculated. Therefore, it may be more promising to measure participants' 'subjective feeling of awareness' in order to compensate for the demerits of the objective measures and to capture the subtle or graded nature of awareness during learning.

The next sections consider the subjective measures of awareness which have attracted researchers' focused attention in the literature since they seemingly capture the more subjective nature of awareness during learning (Wierzchon et al., in press). These subjective measures also meet recent interest in the interplay between cognition and emotion (e.g., Derakhshan & Eysenck, 2010). Since subjective measures require feeling-based judgment in the response, these are employed together with some objective measures such as GJ. We begin consideration of these by reviewing two classical subjective measures: the mere exposure effect and the confidence rating.

The mere exposure effect

The mere exposure effect is that 'repeated, unreinforced exposure results in an increase in positive affect toward a stimulus' (Bornstein, 1989, p.265) and has been extensively investigated and confirmed in various aspects of cognition, such as food or environment preferences (Bornstein, 1989; Zajonc, 2001). Manza, Zizak, and Reber (1998) recommend the mere exposure effect as an alternative test of awareness in implicit learning. The mere exposure effect is argued to be more preferable than standard GJ in that; (1) it does not require recollection of or depend on previous episodes and participants simply rate the emotional preference of a particular sequence. This meets the criteria for implicit cognition developed by Schacter and Graf (1986, p.432) i.e., 'Implicit memory occurs when test performance is facilitated without deliberate or conscious remembering of a study episode'. (2) Since GJ involves awareness of the existence of the underlying rules (i.e. participants are informed about the existence of the rules before GJ), this explicit information on the exsistence of the rules could contaminate GJ that Reber (1967) supposed were a test of implicit knowledge (Manza et al., 1998). That is, such information on the rules, not training on AG, might affect test performance. Conversely, participants in the mere exposure experiment are not informed of the existence of the rules and in this respect, Manza et al. argue, the effects can avoid the problem of explicit contamination.

Initial application to implicit learning was conducted by Gordon and Holyoak (1983). GJs with a three-point confidence rating and six-point Likert rating were administered on the new grammatical and ungrammatical strings of AG. They found that participants could distinguish grammatical from ungrammatical strings and preferred grammatical to ungrammatical strings. The Gordon and Holyoak (1983)

study confounded the scores on the confidence rating with grammaticality, however, (e.g. 1-a very confident ungrammatical vs. 6-a very confident grammatical) and therefore significant differences between grammatical and ungrammatical strings in the scores on the GJ test cannot guarantee the *implicitness* of learning. That is, participants show either differences in confidence or differences in grammaticality. The fact that grammatical strings received higher-preference scores cannot be taken as evidence for implicit learning since there was no independent test of awareness in the Gordon and Holyoak (1983) study.

Manza and Bornstein (1995), on the other hand, did include the recognition test with six-point confidence rating as an independent test for awareness. The experimental groups were asked for GJ whereas the other group was asked for the liking rating after recognition. The results showed that the GJ group considered grammatical strings as more grammatical than the ungrammatical strings whereas the liking-rating groups preferred more grammatical strings than the ungrammatical strings. The recognition performance was above chance in both groups, however, suggesting that the learning processes were in fact 'explicit'. Therefore, contra the authors' assertion that the mere exposure effects can be used for a measurement for awareness, the results of the liking rating did not necessarily capture implicitness of the learning processes either because no independent measures of awareness were used (Gordon & Holyoak, 1983) or because dissociation between explicit and implicit measures or between the explicit measure (recognition) and performance (GJ) was not observed (Manza & Bornstein, 1995). In other words, there was no evidence of the mere exposure effect as the measurement of implicit knowledge and even when this is supposed to be the case the results actually indicate that learning is explicit.

The problem with the mere exposure effects is the lack of the distinction between the two types of. If the mere exposure effect is applied to implicit learning, two effects should be distinguished; (1) the (classic) mere exposure effect: repeated strings (old grammatical strings used in training) promote higher preference, and (2) the structural mere exposure effect: the classic mere exposure effect is generalised to the new grammatical strings that follow the same grammar as the training stimuli (Zizak & Reber, 2004). Of these, only the structural mere exposure effect can be taken as evidence for implicit learning *if it is yoked with* the classic mere exposure effect because implicit learning should involve generalisation or transfer of knowledge acquired during training to different new stimuli that also follow the same grammar.

Only the structural mere exposure effect has been exclusively investigated in the implicit learning literature, however (Gordon & Holyoake, 1983; Manza & Bornstein, 1995; Newell, 2003; Newell & Bright, 2001). These studies involve 'new' grammatical and ungrammatical strings, but not old grammatical training strings, as test stimuli

(Zizak & Reber, 2004, is an exception). For instance, New-ell (2003) found that when training stimulus was presented briefly (100ms) and thus participants' recognition perfor-mance was below chance, no structural mere exposure effect was obtained but the mere exposure effect was ob-served. The latter findings suggest, Newell (2003) claims, that the classical mere exposure effect is not evidence for implicit *learning* since participants simply prefer old train-ing stimuli. Unfortunately, Newell (2003) found these by investigating two types of the mere exposure effect in two separate experiments. Therefore, it is unclear whether the classic mere exposure effect obtained by the old training stimuli are transferred to preference for the ***new*** grammati-cal stimuli as the structural mere exposure effect indicates.

Confidence rating

Another typical subjective measure is confidence rat-ing, which is concerned with how confident participants are about their judgments, as in the case of the mere exposure effects discussed above. Confidence rating has received constant use and still remains extensively employed with other subjective measures (e.g. Brevers et al., 2012; Meal-or & Dienes, 2012a; Schlagbauer, Muller, Zehetleitner, & Geyer, 2012; Wierzchon et al., in press). Tacitly or explic-itly assumed in the confidence rating is the zero-correlation criterion: there should be no correlation between accuracy of performance and confidence, which is evidence for im-plicit learning, given that the confidence rating reflects the degree of awareness (Dienes, 2008; Norman, 2010; Tun-ney, 2005). Some even found that the confidence rating had greater sensitivity to conscious awareness than the verbal report did (Ziori & Dienes, 2006).

Dulany et al. (1984) asked participants to underline the parts of the strings that they thought were grammatical or to cross out the parts of the strings that they thought were ungrammatical. In addition, participants rated their confi-dence on each marked string on a seven-point scale from 'completely uncertain' to 'completely certain' (but the cor-relation between learning and the rating was not analysed). Perruchet and Pacteau (1990) asked participants to rate the degree of recognition of the bigrams (e.g. XX) after training on a six-point scale from 'You are sure that this pair was never present in the strings studied' to 'You are sure that this pair was parts of one or several strings of letters'. Then they divided the scores into recognised vs. unrecognised categories adjusting variance among each participant's data. Vokey and Brooks (1992) also used a six-point rating scale from 'sure to obey the rules' to 'sure not to obey the rules' for GJ, and from "sure similar to at least one training items" to "sure not similar to any training items" for similarity judgment.

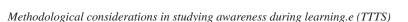Similar six-point scales were also implemented in Shanks et al. (2003), Destrebecquz and Cleermans (2001), and Norman et al. (2006). For instance, Shanks et al. (2003) first asked participants to choose whether the test sequence was old or new, and then participants rated their recognition on a three-point scale on a trial-by-trial basis ('sure', 'fairly sure', and 'guess', a total of six points). Destrebecquz and Cleermans (2001) and Norman et al. (2006) used the fol-lowing scale: from 'I'm certain that this fragment was part of the training sequence' to 'I believe that this fragment was part of the training sequence' to 'I'm certain that this fragment was not part of the training sequence'.

The results of these ratings were sometimes transformed into the binary categories or remained as continuous data. Norman et al. (2007), on the other hand, used the rating with the binary choices from the start. They used 'less/more familiar' for recognition and 'less/more confident' for generation. Yet Dienes and Scott (2005) and Fu et al. (2010) constructed five categories; (1) guess, (2) intuition, (3) pre-existing knowledge, (4) rules, and (5) memory for GJ (Dienes & Scott, 2005) or for sequence generation (Fu et al., 2010).

These authors found positive relations between learn-ing and the degree of awareness on at least some measures of learning and thus provided evidence against learning without awareness, except for Destrebequz and Cleermans (2001), Dienes and Scott (2005), and Fu et al. (2010). Meth-odological issues related to the rating scale, however, urge caution in interpreting these results at face value. First, as reviewed above, there are huge variations in wording in the scale, and particularly problematic is 'guess', which is taken to capture unawareness and is often located on the middle between two extremes (e.g. between 'completely sure to obey the rules' and 'completely sure not to obey the rules'). Dienes (2008) notes that participants differs in their feeling of guessing and they claim to guess even when they know some knowledge. Therefore, Dienes and Scott (2010) suggested that researchers should explain the term *guess* to participants in a clear definition: e.g. '"guessing" means "you know nothing at all—you could just as well have flipped a coin to determine your answer", not just "not very confident"' (p.678).

Second, and in relation to the first issue, it has been ar-gued that participants might be ***reluctant*** to insist on con-fidence when they have 'low' confidence even though they actually know the rules (Perruchet, 2008; Shanks et al., 2003; Wierzchon et al., in press). Again, there is a danger of underestimating awareness.

Third, whether the rating should be dichotomous or con-tinuous is unclear. Currently, evidence for both is available (Dienes, 2008; Tunney, 2005; Tunney & Shanks, 2003). Tunney, by replicating Vokey and Brooks (1992), compared continuous with binary confidence ratings on GJ. He devel-oped a four-point scale; (1) 'yes, confirms to rules—more confident', (2) 'yes, confirms to rules—less confident', (3) 'no, does not confirm to rules—more confident', and (4) 'no, does not confirm to rules—less confident'. In this way, both yes and no response categories became binary (more

or less). In addition, participants were also asked to rate the decision on a 50% to 100% continuous scale where 50% confidence meant a complete guess. The results showed that high confidence in terms of the binary rating increased as the item similarity that was measured on varieties of similarity scales (e.g. associative chunk strength; Perruchet, 1994) increased. The reverse was also the case. In contrast, the continuous measure did not show this trend in many of the similarity measures involved. Therefore, Tunney concluded that the confidence rating with the binary option could capture more subtle states of awareness at the lower level. Similar results using the same measures as Tunney (2005) were obtained in Tunney and Shanks (2003).

On the other hand, Dienes (2008) reported results against the superiority of the binary rating of unpublished ongoing experiments. He compared various confidence ratings including the high/low binary rating and the 50% to 100% continuous rating used in Tunney (2005) and Tunney and Shanks (2003) and obtained null results in terms of ANOVAs. Specific comparison using t-tests revealed, however, that the continuous scale outperformed the binary scale in terms of sensitivity to lower levels of awareness. His conclusion was that no one measure is superior to the other rating measures. As described, sometimes the data on the continuous scale are converted into the binary category (e.g., Perruchet & Pacteau, 1990). This means, however, that there is no independent theoretical motivation to decide the cut-off point (Reber, 1990), which results in arbitrary division.

Fourth, Dienes and colleagues (Dienes & Scott, 2005; Fu et al., 2010) raised the possibility that participants' structural knowledge was unconscious whereas their judgment knowledge was conscious and thus inferring the state of consciousness of the former knowledge from the results of the latter knowledge might be wrong. According to them, judgment knowledge is 'the knowledge directly expressed by a judgment' and structural knowledge is 'the knowledge of the structure of a domain that enabled the judgment, i.e., the basis of (reason for) their judgment' (Fu et al., 2010, p.463). As an analogy, they provided the case of natural language. Native speakers know that a particular sentence is odd (conscious judgment knowledge) but have no idea why this is the case (unconscious structural knowledge). The result is 'intuition'. On the other hand, when both structural and judgment knowledge are unconscious, the result is a 'guess'.

On this basis, they asked participants to choose five categories such as guess, intuition, pre-existing knowledge, rules, and memory in the GJ task. The latter two (rules and memory) indicated the conscious structural rules. Dienes and Scott (2005) found that the number of those which indicated implicit structural knowledge (guess and intuition) outperformed those which indicated explicit counterparts (rules and memory). It was also found that memory and rule together were more related to correct performance than guess and intuition together were but all were significantly beyond baseline performance. These results suggest that participants had conscious and unconscious structural and judgment knowledge. Similar results were replicated in Fu et al., (2010), with the SRT task. What these studies reveal is the importance of dissociation between different types of knowledge though results are still scarce in the literature. In other words, those studies including only the categories *guessing*, *intuition* or *confident* would confound two knowledge types.

In sum, these two traditional subjective measures of awareness, the mere exposure effect and the confidence rating, have raised a number of methodological issues: for the mere exposure effect, the lack of studies which yoked the classic mere exposure effect (likeness for old (training) stimuli) with the structural mere exposure effect (likeness for new (test) stimuli that follow the same rule as the old stimuli); for the confidence rating, shortcomings are more serious as there is ambiguity in wording in the scales, participants might be reluctant to claim their confidence unless they have enough confidence, whether binary or continuous scales should be used is unclear, and structural and judgment knowledge should be separately rated.

The final section deals with post-decision wagering, a new method of measuring awareness in the literature.

Post-decision wagering

Post-decision wagering (Brevers et al., in press; Mealor & Dienes, 2012b; Persaud & McLeod, 2008; Persaud, McLeod, & Cowey, 2007; Wang, Krajbich, Adolphs, & Tsuchiya, 2012; Wierzchon et al., in press) is a measure recently developed as an alternative option to confidence rating. In this task, participants bet real or token money on their decisions and, if their decisions are correct, they get the money. On the other hand, they lose the money if their judgments are incorrect. The rationale behind post-decision wagering is people's tendency to bet a lot of money when they are convinced that their decisions are correct. Therefore, the resulting knowledge is argued to be acquired implicitly when their correct decisions are beyond chance and there is no difference between high- and low-wagering on the correct response.

Persaud et al. (2007) found evidence of implicit AGL. Participants, after GJ of each test string, were asked to bet either one or two pounds on their judgments. The result showed that participants' correct performance was beyond chance (81%), suggesting that they learned the AG. The number of high-wagering participants was not beyond chance, however. This dissociation between learning and wagering was also observed in another group of subjects who betted the token money instead of the real money. Although GJ of the latter group was somewhat less accurate than that of the former group (68%), there were still no differences between high- and low-wagering. These results suggest that post-decision wagering measures the state of

awareness. A similar dissociation between performance and wagering was found in a letter detection task by Persaud and McLeod (2008).

Some advantages of post-decision wagering over the verbal report or the confidence rating are that: (1) instead of forcing participants to introspect their decision, which might cause interference with awareness itself, post-decision wagering is more implicit and less obtrusive, (2) practically speaking, it is more natural and intuitive, or in other words it has much ecological validity, (3) it measures awareness more directly than other measures (other subjective measures measure awareness of awareness, or whether participants are aware that they are aware of something) (Koch & Preuschoff, 2007; Persaud et al., 2007; Wierzchon et al., in press). Koch and Preuschoff (2007) even cite evidence that in the case of the confidence rating a small change in instructional orientation (e.g. whether participants are asked to use all response options equally well) makes the results different (Kolb & Braun, 1995; Morgan, Masson, & Solomon, 1997, see Table 1. *Advantages and Dis- advantages of Experimental Methods in Measuring Aware- ness during Learning*).

A problem specific to post-decision wagering, however, was suggested and tested by Dienes and Seth (2010). The identified problem is risk aversion: when people have low confidence about the correctness of their decision, they might be reluctant to wager much money, even though they have some confidence in their decision. This is because, if their decisions are incorrect, they lose the money betted. Therefore, apparent dissociations between performance and wagering do not necessarily show unawareness but indicate the effects of risk management. In order to clarify these relations, Dienes and Seth (2010) directly compared confidence rating with a binary option (guess or sure) with post-decision wagering including a risk aversion task. In the risk aversion task, participants were asked two questions: (1) 'If there was a lottery for a 10 € prize, which would be given to one of the 10 ticket holders, how much would you pay for a ticket?', and (2) 'If the prize were 100€, which would be given to one of the 10 ticket holders, how much would you pay for a ticket?' (pp.676-677).

Sweets rather than actual money were used in Dines and Seth's experiments (2010). The results revealed that participants showed higher confidence and higher-vetting on the correct decision. These were significantly above chance but did not differ from each other. There were differences, however, between the confidence rating and the post-decision wagering on the proportion of the low confidence response (wagering > the confidence rating), suggesting that participants were more reluctant to express confidence by wagering than by the confidence rating. Furthermore, there were significant negative correlations between risk aversion and wagering, that is, as risk aversion became high, wagering became low. Conversely, no significant correlation was observed between risk aversion and confidence rating.

These results suggest that confidence rating with the binary option is a more sensitive measure of awareness than post-decision wagering. Interestingly, in their experiment 2, instead of the post-decision wagering, they administered a no-loss gambling task where participants had to decide either to stay with their decision or bet on the card that gave sweets on a 50 % probability (see also Mealor & Dienes, 2012b). Importantly, participants did not need to lose the sweets when their decisions were wrong. The results showed that the amount of betting on the card was significantly lower than either the confidence rating or the wagering and that no correlation was found between the no-loss gambling and the risk aversion. This suggests that if the appropriate controls (e.g. risk aversion) are applied, gambling methods can be a possible measure of awareness like confidence rating.

Taken together, subjective (and emotional) measures have one important limitation: the subjective nature of the awareness measures itself brings aversion activities into and thus contaminates the experiment: participants do not express their subjective and emotional feelings such as likeness, confidence, or motivation for vetting unless they have sufficient confidence in their decisions. Therefore, subjective measures have a great risk of underestimating awareness during learning. Table 1 provides a summary of the advantages and disadvantages of each method discussed in this paper.

Conclusion

By now, it is clear that separating implicit from explicit learning solely based on learning instructions or populations is unreliable. There is no guarantee that participants will follow the instructions whether the same participant engages in a series of implicit and explicit learning tasks or different participants engage in implicit or explicit learning tasks. In addition, amnesics differ from their normal peers even in performance in implicit learning tasks and thus the results of experiments recruiting the former population are not generalisable to the latter population. Therefore, it might be preferable to carry out a single learning task with some measures of awareness and classify participants into aware/unaware groups according to those measures.

How awareness during learning should be measured remains unclear, however. Any measures of awareness face difficulty in meeting Shanks and St. John's (1994) information and sensitivity criteria. This is because of methodological problems: some of them such as participants' reluctance to claim confidence when they are less confident even though there is partial awareness are broadly applied to many measures. Others are more peculiar to the particular measure: whether dichotomous or continuous rating should be employed (confidence rating) or whether participants try to avoid money loss (post-decision wagering). Therefore, some even argue that the verbal report is still the

Methods	Operationalisation	Examples	Advantages	Disadvantages
Employing different populations (e.g., .. amnesiacs)	Comparing different populations on implicit learning performance	Knowlton & Squire (1996)	-	Amnesics show different (impaired) implicit learning performance from the normal and thus the results are not generalizisable.
Employing different learning instructions	(a) Implicit learning instruction: mmemorizisation of training stimuli (b) Explicit learning instruction: intentional rule search on training stimuli	Reber (1967) Reber et al. (1991)	-	(a) Results comes from task difference (between-participants design) (b) Learning tasks are contaminated (within-participants design) (c) Participants do not simply follow instructions
The verbal report	(a) Ask participants to disclose contents of awareness after experiments (retrospective) (b) Ask participants to disclose contents of awareness during experiments (concurrent)	Reber (1967)	Verbal report can measure higher-order (access) consciousness	(a) Participants may forget the contents of awareness after experiments (retrospective) (b) Verbal report may interfere with ongoing learning process (concurrent) (c) Verbal report cannot measure phenomenal awareness

Objective measures

Methods	Operationalisation	Examples	Advantages	Disadvantages
(1) Dichotomous GJ, recognition, generation	(a) Ask participants to judge (b) whether a new stimuli is grammatical (c) whether a new stimuli is what they saw during training (d) where the light appears on next given fragmentary sequences of a stimuli	Reber (1967) Destrebecqz & Cleermans (2001)	(a) Objective measures are more objective than the verbal report because participants are simply required to categorise stimuli into two categories (e.g., *seen* and *unseen*) (b) Objective measures are more sensitive than the verbal report since participants can respond on the basis of phenomenal awareness.	(a) GJ is a measure of explicit knowledge instead of implicit knowledge (b) The dichotomous discrimination without scales is too coarse- grained to capture the subtle nature of participants' awareness

(2) strategic control	Give two instructions: (a) inclusion: ask participants to include the training stimuli that they had saw (b) exclusion: ask participants to exclude the training stimuli that they had saw *Exclusion errors reflect the effects of unconscious processes *Unconscious knowledge = inclusion - exclusion	Dienes, et al. (1995)	(a) The problem of the task differences can be avoided because the PDP in the implicit learning literature usually employs instructional differences for the same task instead of task differences (b) The problem of task contamination can be resolved by statistical calculation of inclusion and exclusion performances	(a) Exclusion errors do not reflect effects of unconscious processes but result from external factors such as motivation (b) Exclusion instruction is too difficult to follow (c) When there are no exclusion errors, the magnitude of the unconscious processes cannot be estimated

Subjective measures

(1) The mere exposure effects	Ask participants to judge whether they like a stimulus	Gordon & Holyoak (1983)	(a) The effect does not require recollection of or depend on previous episodes and participants simply rate the emotional preference of a particular sequence (b) Participants are not informed of the existence of the rules and in this respect, the effects can avoid the problem of explicit contamination	Experiments often investigate either (1) the classical mere exposure effect (likeliness for old (training) stimuli or (2) the structural mere exposure effect (likeliness for new (test) stimuli, but not both.
(2) The confidence rating	(a) Ask participants to rate confidence of their response on binary or continuous scales (b) Ask participants about their degree of confidence by selecting one from several response categories (e.g., guess, intuition)	Perruchet & Pacteau (1990) Dienes & Scott (2005)	The measure captures better the subjective and graded nature of awareness during learning	(a) There is variability or ambiguity in wording of response options (b) Participants might be reluctant to claim confidence unless they have enough confidence (c) Whether the rating should be dichotomous or continuous is unclear (d) Structural knowledge and judgment knowledge might be confounded

(3) The post-decision wagering	Ask participants to bet money on their decision; and if their decision is correct, they get the money	Persaud, et al. (2007)	(a) Instead of forcing participants to introspect their decision, which might cause interference with awareness itself, the post-decision wagering is more implicit and less obtrusive (b) Practically speaking, the method is more natural and intuitive, or, in other words, the method has much ecological validity (c) The method measures awareness more directly than other measures (other subjective measures measure awareness of awareness, or whether participants are aware of that they are aware of something)	When people have low confidence about the correctness of their decision, they might be reluctant to wager much money, even though they have some confidence in their decision

gold standard as regards measurement of awareness (Runger & Frensch, 2010; Weiskranz, 1997). Overall, as Mealor and Dienes (2012b, p.236) claim, one might conclude

> But one can never be sure that any method always picks out just unconscious knowledge….Any measuring method can be criticised for the mere possibility it might sometimes get it wrong: the acid test is if in practice it gets it right often enough that it participates in theory-driven research, providing itself by the theories it can corroborate.

According to Mealor and Dienes (2010b), verbal reports may be a reliable measure in that access to verbalisable consciousness (theoretically-driven; Block, 2005, 2007) by definition can be measured only by some kind of verbal report.

After all, capturing the null effect of awareness is difficult, if not impossible. According to Erdelyi's (1986) experimental indeterminacy, implicit knowledge can be demonstrated only in the condition $\alpha = 0$ *and* $\varepsilon > 0$, where α refers to conscious knowledge and ε refers to unconscious knowledge. Reber (1989b) argues that this over-strict condition cannot be met. Reber himself (1989b) concedes that participants in his implicit learning experiments developed some awareness of the underlying structures: they were not totally unaware of the grammar. This is also reflected in his definitions of im-

plicit learning: 'largely independently' from awareness. It is, then, safer to conclude that there are still controversies over the possibility of learning without awareness (Baars, 2002).

One direction for current and future research is to clarify limitations inherent in a particular measure by employing theoretically-driven ***multiple*** subjective measures of awareness and to research differential sensitivities to ***graded*** awareness during learning. The key to doing so is interdisciplinary sensitivity. One example is Wierzchon et al. (in press), who employed feelings of warmth (cold, chilly, warm, and hot), which had originally developed as a measure of *intuition* (Metcalfe,1986), as a subjective measure of awareness to capture subtle awareness (Kouider, de Gardelle, Sackur, & Dupoux, 2010). Much research along this line is clearly needed.

References

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*, 47-52.

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, *39*, 585-609.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9, 46-52.*

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioural and Brain Sciences*, *30*, 481-548.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, *106*, 265-289.

Brevers, D., Cleeremans, A., Bechara, A., Greisen, M., Kornreich, C., Verbanck, P., & Noel, X. (in press). Impaired metacognitive capacities in individuals with problem gambling. *Journal of Gambling Studies*.

Brody, N. (1989). Unconscious learning of rules: Comment on Reber's analysis of implicit learning. *Journal of Experimental Psychology: General*, *118*, 236-238.

Brown, J., Aczel, B., Jimenez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *Quarterly Journal of Experimental Psychology*, *63*, 1789-1812.

Channon, S., Shanks, D. R., Johnstone, T., Vakil, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia?: Rule abstraction and item familiarity in artificial grammar learning. *Meuropsychologia*, *40*, 2185-2197.

Curran, T. (1997). Higher-order associative learning in amnesia: Evidence from the serial reaction time task. *Journal of Cognitive Neuroscience*, *9*, 522-533.

Debner, J. A. & Jacoby, L. L. (1994). Unconscious perception: Attention, awareness, and control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 304-317.

Della Sala, S., Cowan, N., Beschin, N., & Perini, M. (2005). Just lying there, remembering: Improving recall of prose in amnesic patients with mild cognitive impairment by minimizing interference. *Memory*, *13*, 435-440.

Derakhshan N. & Eysenck, M. (Eds.). (2010). *Emotional States, Attention, and Working Memory*: *A Special Issue of Cognition & Emotion*. New York: Psychology Press.

Destrebecqz, A., & Cleermans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*, 343-350.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research*, *168*, 49-64.

Dienes, Z., Altman, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammar is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1322-1338.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, *69*, 338-351.

Dienes, Z., & Scott, R. (2010). Gambling on the unconscious: A comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Consciousness and Cognition*, *19*, 674-681.

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541-555.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305-352.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141-172.

Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip, & H. Reinders (Eds.), *Implicit and explicit knowledge of second language learning, testing, and teaching* (pp. 31-64). Bristol: Multilingual Matters.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005). *Studies in Second Language Acquisition*, *29*, 119-126.

Ellis, R., Leowen, S., Elder, C., Erlman, R., Philip, J., & Renders, H. (Eds.). (2009). *Implicit and explicit knowledge of second language learning, testing, and teaching*. Bristol: Multilingual Matters.

Erdelyi, M. H. (1986). Experimental indeterminacies in the dissociation paradigm of subliminal perception: Comment on Holender (1986). *Behavioural and Brain Sciences*, *9*, 30-31.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Boston, MA: MIT Press.

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, *67*, 279-300.

Fisk, G. D., & Haase, S. J. (2006). Exclusion failure does not demonstrate unconscious perception II: Evidence from a forced-choice exclusion task. *Vision Research*, *46*, 4244-4251.

Frensch, P. A. (1998). One concept, multiple meanings: On how to define the concept of implicit learning. In M. A. Stadler & P. A. Frensch, (Eds.), *Handbook of implicit learning* (pp.47-104). Thousand Oaks, CA: Sage.

Fu, Q., Dienes, Z., & Fu, X. (2010). Can unconscious knowledge allow control in sequence learning? *Consciousness and Cognition*, *19*, 462-474.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33*, 34-54.

Gordon, P. C., & Holyoak, K. J. (1983). Implicit learning and generalization of the 'mere exposure' effect. *Journal of Personality and Social Psychology*, *45*, 492-500.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioural and Brain Sciences*, *9*, 1-66.

Hutchinson, K. A., Neely, J. H., Neill, W. T., & Walker, P. B. (2004). Is unconscious identity priming lexical or sublexical? *Consciousness and Cognition*, *13*, 512-538.

Jacoby, L. L. (1991). A process-dissociation framework: Separating automatic from intentional use of memory. *Journal of Memory and Language*, *30*, 513-541.

Jourdenais, R. (2001). Cognition, instruction, and protocol analysis. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 354-375). New York, NY: Cambridge University Press.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of classification learning and explicit memory for specific instances. *Psychological Science*, *3*, 172-179.

Knowlton, B. J., & Squire, L. R. (1994). Information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 79-91.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 169-181.

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Nature Neuroscience*, *10*, 140-141.

Kolb, F. C., & Braun, J. (1995). Blindsight in normal observers. *Nature*, *377*, 336-338.

Kouider, S., de Gardelle, V., Sackur, J., & Dupoux, E. (2010). How Rich is Consciousness? The Partial Awareness Hypothesis. *Trends in Cognitive Sciences*, 14, 301-207.

Manza, L., & Bornstein, R. F. (1995). Affective discrimination and the implicit process. *Consciousness and Cognition*, *4*, 399-409.

Manza, L., Zizak, D., & Reber, A. S. (1998). Artificial grammar learning and the mere exposure effect. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp.201-222). Thousand Oaks, CA: Sage.

Mealor, A. D., & Dienes, Z. (2012a). Conscious and unconscious thought in artificial grammar learning. *Consciousness and Cognition*, *21*, 865-874.

Mealor, A. D., & Dienes, Z. (2012b). No-loss gambling shows the speed of the unconscious. *Consciousness and Cognition*, *21*, 228-237,

Metcalfe, J. (1986). Feeling of knowing in memory and problem-solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *12*, 288-294.

Morgan, M. J., Masson, A. J. S., & Solomon, J. A. (1997). Blindsight in normal Subjects? *Nature*, *385*, 401-402.

Nakamura, D. (accepted). Methodological considerations in studying awareness during learning: Part 2. Second language acquisition. Polish Psychological Bulletin.

Newell, B. R. (2003). The subliminal mere exposure effect does not generalize to structurally related stimuli. *Canadian Journal of Experimental Psychology*, *57*, 61-68.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirement of learning: Evidence from performance measures. *Cognitive Psychology*, *19*, 1-32.

Norman, E. (2010). 'The unconscious in current psychology'. *European Psychologist*, *15*, 193-201.

Norman, E., Price, M. C., & Duff, S. C. (2006). Fringe consciousness in sequence learning: The influence of individual differences. *Consciousness and Cognition*, *15*, 723-760.

Norman, E., Price, M. C., Duff, S. C., & Mentzoni, R. A. (2007). Gradations of awareness in a modified sequence learning task. *Consciousness and Cognition*, *16*, 809-837.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 223-228.

Perruchet, P. (2008). Implicit learning. In H. L. Roediger, III (Ed.), *Learning and memory, volume 2. Cognitive psychology of memory* (pp.597-621). Oxford: Elsevier.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264-275.

Persaud, N., & McLeod, P. (2008). Wagering demonstrates subconscious processing in a binary exclusion task. *Consciousness and Cognition*, *17*, 565-575.

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

Poznanski, Y., & Tzelgov, J. (2010). Mode of knowledge acquisition and retrieval in artificial grammar learning. *Quarterly Journal of Experimental Psychology*, *63*, 1495-1515.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behaviour*, *6*, 855-863.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology; Human Learning and Memory*, *2*, 88-94.

Reber, A. S. (1989a). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219-235.

Reber, A. S. (1989b). More thoughts on the unconscious: Reply to Brody and to Lewicki and Hill. *Journal of Experimental Psychology: General*, *118*, 242-244.

Reber, A. S. (1990). On the primacy of the implicit: Comment on Perruchet and Pacteau (1990). *Journal of Experimental Psychology: General*, *119*, 340-342.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York, NY: Oxford University Press.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, *17*, 888-896.

Richard, M. V., Clegg, B. A., & Segar, C. A. (2009). Implicit motor sequence learning is not represented purely in response locations. *Quarterly Journal of Experimental Psychology*, *62*, 1516-1522.

Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, *47*, 45-99.

Robinson, P. (2002). Individual differences in intelligence, aptitude, and working memory during adult incidental second language learning: A replication and extension of Reber, Walkenfeld, and Harnstadt (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211-266). Amsterdam: John Benjamins.

Robinson, P. (2005). Cognitive abilities, chunk-strength and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, *15*, 413-438.

Robinson, P. (2009). Replicating and generalizing findings from cognitive psychology to second language acquisition: conceptual and methodological issues. Paper presented at the invited symposium on 'Replication Research in Applied Linguistics and SLA' convened by Graeme Porte, at the American Association of Applied Linguistics Conference (AAAL), March, 2009, Denver, USA.

Runger, D., & Frensch, P. A. (2010). Defining consciousness in the context of incidental sequence learning: Theoretical considerations and empirical implications. *Psychological Research*, *74*, 121-137.

Schacter, D. L., & Graf, P. (1986). Effects of elaborate processing on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 432-444.

Schlagbauer, B., Muller, H. J., Zehetleitner, M., & Geyer, T. (2012). Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks. *Journal of Vision*, *12*, 1-12.

Segar, C. A., Parabhakaran, V., Poldrack, R. A., & Gabrieli, J. D. E. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychbiology*, *28*, 283-292.

Seth, A. K., Dienes, Z., Cleermans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*, 314-321.

Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202-220). London: Sage.

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*, 369-382.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences*, *17*, 367-447.

Shanks, D. R., Wilkinson, L., & Channon, S. (2003). Relationship between priming and recognition in deterministic and probabilistic sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 248-261.

Snodgrass, M. (2002). Disambiguating consciousness and unconsciousness influences: Do exclusion paradigms demonstrate unconscious perception? *American Journal of Psychology*, *115*, 545-579.

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: Replicable and qualitatively different unconscious perceptual effects. *Cognition*, *101*, 43-79.

Tunney, R. J. (2005). Sources of confidence judgments in implicit cognition. *Psychonomic Bulletin & Review*, *12*, 367-373.

Tunney, R. J., & Shanks, D. R. (2003). Does opposition logic provide evidence for conscious and unconscious processes in artificial grammar learning? *Consciousness and Cognition*, *12*, 201-218.

Visser, T. A. W., & Merikle, P. M. (1999). *Consciousness and Cognition*, *8*, 94-113.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 328-344.

Wang, S., Krajbich, I., Adolphs, R., & Tsuchiya, N. (2012). The role of risk aversion in non-conscious decision making. *Frontiers in Psychology*, *3*, 1-17.

Weiskranz, L. (1997). *Consciousness lost and found: A neuropsychological exploration*. New York, NY: Oxford University Press.

Wierzchon, M., Asanowicz, D., Paulewicz, B., & Cleermans, A. (in press). Subjective measures of consciousness in artificial grammar learning task. *Consciousness and Cognition*.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 354-369.

Yonelinas, A. P., & Jacoby, L. L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Memory and Cognition*, *40*, 663-680.

Ziori, E., & Dienes, Z. (2006). Subjective measures of unconscious knowledge of concepts. *Mind & Society*, *5*, 105-122.