

Well past midnight: Calling time on implicit invariant learning?

Ben R. Newell and James E.H. Bright

University of New South Wales, Sydney, Australia

Three experiments are reported that examine the nature of knowledge underlying performance in the invariant learning task. Previous research (Bright & Burton, 1994; McGeorge & Burton, 1990) has supported an account of performance based on the implicit abstraction and application of a rule pertaining to the invariant feature. In contrast, we found effects in both the digit and clock invariant tasks that are difficult to explain solely in terms of subjects acquiring the experimenters' rule. In all three experiments, manipulation of test item properties that are independent of the invariant feature led to a detriment in performance that is not predicted by an account based on the experimenters' rules. Furthermore, the use of an on-line measure of awareness (confidence ratings) provided some evidence that performance is mediated by low confidence explicit knowledge.

Early accounts of implicit learning—a process of unintentionally acquiring a sensitivity to a set of stimuli—posited a powerful, unconscious mechanism capable of abstracting rules that describe the underlying structure of a stimulus environment (e.g., Reber, 1967, 1976). Under such an account, performance improvements are attributable to the application of the acquired rules to novel stimuli. Subsequent accounts (e.g., Brooks, 1978; Perruchet & Gallego, 1997) state that implicit learning is not characterised by rule acquisition but rather by the storage of discrete instances in memory. According to this view, improvements in performance result from a similarity-match between novel stimuli and stored instances, or fragments thereof (e.g., Perruchet & Pacteau, 1990), of previously encountered stimuli.

Requests for reprints should be addressed to B. Newell, Department of Psychology, University College, London, Gower Street, London, WC1E 6BT, UK. Email: b.newell@ucl.ac.uk

Ben Newell was supported by a Commonwealth Scholarship. Jim Bright was supported by an Australian Research Council Small Grant. We would like to thank Peter Lovibond, Fred Westbrook, Axel Cleeremans, Pierre Perruchet, and one anonymous reviewer for their helpful comments on an earlier draft of this paper. We also thank Sally Andrews for assisting in collecting the data for Experiment 2.

The relative merits of these two accounts have been extensively researched over the last 20 years, and it has become clear that an explanation of implicit learning based solely on implicit rule abstraction has become increasingly difficult to justify (Gomez, 1997; Gomez & Schvaneveldt, 1994; Shanks & Johnstone, 1998; Vokey & Brooks, 1992).

One paradigm that reflects this on-going debate is the invariant learning tasks of Burton and colleagues (Bright & Burton, 1994; Cock, Berry, & Gaffan, 1994; McGeorge & Burton, 1990; Wright & Burton, 1995). In one version of this paradigm (the Bright and Burton clocks task) the case for an implicit rule-based account is still strongly argued (Bright & Burton, 1994, 1998), whereas in other versions (the McGeorge & Burton 3s task, 1990) existing studies leave open the possibility of similarity-based processing (Cock et al., 1994; Wright & Burton, 1995).

In the experiments presented here we examine two demonstrations of invariant learning in an attempt to determine the nature of the representations underlying subjects' performance. In particular we aim to assess the capability of an implicit rule-based account to explain the learning effects observed in the invariant paradigm.

Implicit invariant learning is said to have occurred when a person demonstrates an acquired sensitivity to an invariant characteristic of a set of stimuli, in the absence of verbal knowledge of that characteristic. In the original demonstration, McGeorge and Burton (1990) showed subjects 30 four-digit number strings in an incidental learning task. All the strings presented to the subjects contained the digit 3, though this invariant feature was not brought to the attention of the subjects. In a subsequent surprise test phase, subjects were shown 10 pairs of novel four-digit strings, where one string in each pair contained a 3 (the *positive*) and the other did not (the *negative*). Subjects were told (falsely) that they had seen one and only one of the strings in the previous learning phase, and were instructed to choose the one they thought they had seen before. The reliable finding was that subjects chose more positives than would be expected by chance (typically 6 or 7 out of 10). This preference persisted when the surface representation of strings was changed from numeric (3457) during learning, to word (three five two one) at test.

Bright and Burton (1994) extended the original finding to the learning of a more complex invariant feature. Bright and Burton showed subjects 30 schematic clock faces that all depicted times between 6 and 12 o'clock. Thus, the position of the hour hand was the invariant feature. At test subjects were shown 10 novel pairs of clocks with one showing a time inside the invariant range (*positive*) and one 6 hours away from that time (*negative*) (e.g., 8:30 and 2:30). Once again, the robust finding was that subjects showed a selection preference for the positives. A change in surface representation from analogue during learning to digital at test had no effect on subjects' preferences.

The surprising finding in both these demonstrations was that subjects performed above chance at test, in spite of being unable to verbalise any knowledge pertaining to the invariant feature. Indeed many subjects expressed surprise when told about the nature of the underlying invariant rule.

McGeorge and Burton (1990) and Bright and Burton (1994, 1998) argued that the simplest explanation of these findings was that subjects were implicitly learning the critical invariant rule.

However, as a number of investigators have pointed out (Berry & Cock, 1998; Churchill & Gilmore, 1998; Cock et al., 1994; Wright & Burton, 1995; Wright & Whittlesea, 1998), rather than responding on the basis of the information that the experimenter *assumes* the subject has acquired, the subject may be utilising other correlated information to make selections based on the similarity (Cock et al., 1994) or unfamiliarity (Churchill & Gilmore, 1998; Wright & Burton, 1995) of test items in comparison with learning items.

Wright and Burton (1995) suggested that the correlated information used by subjects in the ‘3s’ task was the presence of distinctive features—repetitions—in the strings. An analysis of the typical learning strings used in the McGeorge and Burton experiment showed that constraining all strings to contain a ‘3’ reduced the probability of repeated digits (e.g., 5772 or 7572) occurring. They noted that the probability of repeated digits appearing in a positive (string containing the invariant 3) was 0.34 compared to the probability of 0.59 of repeated digits appearing in a negative (string without a 3). Wright and Burton’s hypothesis was that subjects score at an above-chance level not because the positives conform to an implicitly learnt rule but because it is easy to reject distinctive items (i.e., those with repetitions) seen at test. Their results indicated that 64% of test decisions could be classified as involving rejection of the distinctive item, compared with 59% of occasions in which the positive was selected. Wright and Burton concluded that subjects’ behaviour was better classified as ‘rejecting repetitions’ than ‘accepting positives’ (1995, p. 794).

This explanation of subjects’ behaviour accounts well for the simple version of the invariant 3 task—that is when both learning and test items are represented in the same surface format (i.e., digits). However, in the original demonstration, McGeorge and Burton (1990) showed that subjects’ preferences for items containing the 3 persisted when the surface representation was changed from digits at learning to words at test. To date, the rejection strategy of Wright and Burton (1995) has not been tested to see if it can account for subjects’ behaviour in this cross-format transfer condition. It is possible that the effects observed by Wright and Burton (1995) were based on perceptual recognition where there is an advantage for items that bear the same perceptual characteristics during both the learning and test phases. The importance of perceptual characteristics has been demonstrated in the implicit memory literature where a number of studies have shown that a change in surface representation between learning and test

results in reduced priming effects (e.g., Blaxton, 1989; Jacoby & Hayman, 1987).

In Experiment 1 we test the rejection strategy of Wright and Burton (1995) and examine whether subjects' tendency to reject items containing repeated digits interacts with a change in the perceptual format of items between the learning and test phase. If an interaction were obtained this would suggest that the rejection strategy proposed by Wright and Burton is not a generic mechanism underlying performance in the invariant 3s task. An interaction would suggest that the tendency to reject distinctive items is observed only when the perceptual format of learning and test items is the same. In contrast, the failure to find an interaction would demonstrate that the rejection strategy operates under both same and cross-format transfer conditions.

EXPERIMENT 1

Method

Participants. Thirty undergraduate students from the University of New South Wales participated in the experiment in return for course credit. All were aged between 18 and 30 years. All were naïve with respect to the invariant learning procedure.

Materials. Learning sets were generated by selecting randomly 30 number strings from a database containing all 2048 possible four-digit numbers constructed from the digits 1 to 9 and containing a single digit 3. Test sets were generated from four databases: one containing all possible four digit numbers with a single "3" and with no repetitions, one with a single "3" and with repetitions, one with no "3" and no repetitions, and one with no "3" and with repetitions. Following Wright and Burton (1995) quadruples, triples, and double doubles were removed from the databases leaving only doubles to serve as distinctive items. Items containing a 3 were termed positives and those without a 3 negatives. Test sets contained three types of pairs: those biasing subjects *towards* positive selection, those biasing subjects *against* positive selection, and those that were *neutral* with respect to positive selection. That is:

- (1) the negative contained a double but positive did not—this should bias subjects *towards* selection of the positive (e.g., 5772 vs 2367)
- (2) the positive contained a double but negative did not—this should bias subjects *against* selection of the positive (e.g., 3525 vs 2167)
- (3) that neither the positive nor the negative contained a double—this should be *neutral* with respect to selection of the positive (e.g., 3428 vs 7865).

Each subject received twelve pairs, four from each of the three pair types. All learning items were printed in numeric format (1357), test items were printed in either numeric or word format (one three five seven).

Design and procedure. A mixed design was used. The within-subjects factor was direction of bias, this had three levels: towards, neutral, against. The between-subjects factor was test format, with two levels: same or changed. Subjects were randomly assigned to one of the two groups. The “same” group learned and were tested on numeric items; the “changed” group learned on numeric items and were tested on word items. The dependent variable was the number of positives selected at test.

The experiment consisted of three phases: learning, an unexpected test and post-test questioning. Subjects were shown 30 four-digit numbers containing the digit 3 and were asked to perform an addition and comparison task on each number. This involved adding together the first two digits and comparing the sum with the total of the second two digits. After completing this task, learning materials were removed from sight and subjects were presented with 12 test pairs in numeric (same group) or word (changed group) format and were told (falsely) that they had seen one item in each pair before. The left/right position of the positive and negative and the order of presentation of the three pair types were counterbalanced. Subjects were asked to indicate which item they thought they had seen before and to guess if they were unsure. On completion of the test phase a questionnaire was presented to the subjects (see the Appendix). On completion of the questionnaire, subjects were fully debriefed.

Results

The mean number of positives selected was 6.13 (out of a maximum of 12), $SD = 1.5$. A single sample t -test showed that this was not significantly different from chance performance of 6, [$t(29) = 0.50, p > .1$]. A 2 (Test Format: same, changed) \times 3 (Direction of Bias: towards, neutral, against) mixed model analysis of variance was conducted. This revealed a significant main effect of direction of bias, [$F(2, 56) = 4.62, p < .05$], and no significant effect of test format, [$F(1, 28) = 1.85, p > .05$]. The interaction between test format and direction of bias was not significant, [$F(2, 56) = 1.13, p > .05$]. The mean numbers of positives selected in each cell of the 2 \times 3 design are presented in Table 1.

TABLE 1
Mean selection of positives (out of four) and standard deviations by condition (Experiment 1)

<i>Test format</i>	<i>Towards</i>	<i>Neutral</i>	<i>Against</i>
Same (digit)	2.73 (1.22)	1.93 (1.10)	2.07 (1.10)
Changed (word)	2.53 (1.19)	2.13 (0.99)	1.33 (0.97)

In order to be confident that subjects' tendency to reject items containing repeated digits extended to the changed format condition we conducted a one-way (direction of bias) repeated measures ANOVA on the data from this condition. A significant main effect, [$F(2, 28) = 4.42, p < .05$], was found demonstrating that subjects do tend to reject distinctive items when the perceptual format is changed between learning and test phases. Planned contrasts conducted on combined data from the same and changed format conditions showed a significant difference between the *towards* and *against* condition, [$F(1, 28) = 9.18, MSe = 1.42, p < .05, (Fc = 6.48)$]. Contrasts between *towards* and *neutral* and *against* and *neutral* were not significant. Mean selection scores were then compared against chance performance (2 for each condition). For the *towards* condition, more positives were selected than would be expected by chance, [$t(29) = 2.92, p < .01$]; for the *neutral* and *against* conditions selection was at chance, [$t(29) = .18, p > .1; t(29) = 1.51, p > .1$, respectively].

Following Wright and Burton (1995), the mean selection scores in each condition were converted to percentages to determine the proportion of responses that could be classified as involving rejections. In the *towards* condition items containing doubles were rejected on 65% of occasions. In the *against* condition 42% of items containing doubles were selected leaving 58% rejected. This means that over these two conditions 61% of decisions were classified as involving rejection of items containing repeated numbers. This compares to the 54% of decisions involving selection of the items containing the invariant 3.

Test of explicit knowledge. All subjects ($N = 30$) reported that some of their responses at test were guesses. When questioned about the strategy they employed when guessing, seven could provide no justification for their guesses and a further thirteen mentioned memory or familiarity without elaborating on the way in which it was used. However, four subjects reported relatively detailed strategies including: rejecting items with totals that equalled each other; choosing items with totals closer together (e.g., 7673 over 2198); and choosing the item in which the left pair total was larger than the right pair due to a perceived preponderance of these items in the learning set. Analysis of the learning and test sets used by these subjects revealed that performance predicted by the use of these strategies approximately matched actual performance in all cases. This finding indicates that these subjects were using explicit veridical knowledge and not merely providing post hoc explanations of their behaviour. A further six subjects mentioned grounding their responses on double digits (by double they referred only to contiguous doubles, e.g., 77 or seven seven). These subjects' responses were further analysed to see if they conformed to their explicit knowledge. We found that when the contiguous double was in the negative item (*towards* condition) positives were selected on 83% of occasions; in comparison when the contiguous double was in the positive item (*against*

condition), selection of the positive was at 33%. This suggests that these subjects were responding to some extent on the basis of the presence or absence of double digits. In terms of positive selection these subjects did not appear to perform any better than subjects who did not rely on explicit knowledge of double digits: 6.6 (SD 1.49) compared to 5.5 (SD 1.64). Neither group performed significantly differently from chance ($p > .1$).

Finally, explicit knowledge of the invariant 3 was examined. When asked to circle the number that they thought had appeared most frequently in the learning items, eight subjects circled the number 3. This is more than the 3.3 subjects predicted by chance. The mean positive selection for these subjects was 6.0 (chance performance) and marginally lower than the mean for the whole group (6.13). This finding suggests that knowledge of the 3 was not used in the test phase but may have been cued by the questionnaire. Many subjects expressed surprise when told about the nature of the experiment during debriefing.

Discussion

Experiment 1 has demonstrated that subjects' tendency to reject items containing repeated digits does not interact with a change in the perceptual format between learning and test items. Prior to Experiment 1, the rejection strategy had not been tested in cross-format transfer conditions, leaving open the possibility that the effects observed by Wright and Burton were due to perceptual-based recognition. We can now be confident that the rejection strategy can account for both same-format and cross-format versions of the invariant digit task.

Experiment 1 has provided further, stronger evidence that subjects in the invariant digit task do not learn the rule that it was originally supposed subjects learnt in this task (i.e., "Select items containing a 3"). The question of what knowledge subjects are acquiring remains open to debate. It is possible that subjects are applying an explicit (for some subjects at least) "reject repetitions" rule, or that the rejection is part of a more general similarity-based mechanism in which test strings are compared (implicitly or explicitly) with memories of learning strings. We will return to this debate in the General Discussion and question whether it is possible (within the invariant task) to distinguish between rule- and similarity-based behaviour.

Our finding of further evidence that subjects are not learning the experimenters' rules in the 3s task leads us to question whether performance in the logically related clocks task (Bright & Burton, 1994, 1998) is based solely on the experimenters' rule pertaining to the invariant time range. It is possible that the use of information correlated with the rule that we and others (e.g., Cock et al., 1994; Wright & Burton, 1995) have observed in the digit task may be due to the nature of the numeric stimuli used. It is plausible that certain structural characteristics of the number strings lend themselves to this type of processing. A clearer indication of a generic mechanism underlying invariant learning would

come from finding that the observed effects generalise to the more complex clock stimuli of Bright and Burton (1994, 1998).

EXPERIMENT 2

A logical way to look for similar effects would be to test for a rejection strategy using the clock stimuli. However, the nature of the stimuli make it extremely difficult to design a “rejection” test set. In our minds, there seems to be no correlated feature (such as repetitions) that varies with the invariant time range. Likewise, manipulation of the specific similarity of learning and test items is also problematic due to the way in which the clock stimuli are constructed. For this reason, we manipulated other properties of the test pairs that we hypothesised would have an effect on subjects’ decisions at test. Bright and Burton (1994, 1998) argued that performance in the clock task was indicative of implicit rule abstraction. If subjects are abstracting a rule during learning then presumably it would be of the form “favour times between 6 and 12” (Bright & Burton, 1994, p. 81). If we accept this, then provided one time in each pair is between 6 and 12 and the other time is outside this range we would expect to observe the normal selection preference. Alternatively, subjects may be engaging in a comparison process similar to the one that may underlie the rejection strategy. If subjects are making comparisons of test items with learning items then reducing the similarity between learning and test items should be detrimental to performance. In Experiment 2 we contrast performance on test pairs with differing degrees of intra-test pair similarity. Test pairs with a low similarity between times (those separated by 6 hours, e.g., 8:30 vs 2:30) are compared with pairs with a medium similarity (those separated by 3 hours, e.g., 4:30 vs 7:30) and with pairs with a high similarity (those separated by 1 hour, e.g., 5:10 vs 6:10). If performance is based on the “6 to 12” rule we expect no differences between the pair types. However, if subjects are engaged in similarity-based processing, we expect selection of the invariant in the 1-hour pairs to be lower than in the 6-hour pairs.

Any observed differences between the pair types could be due to subjects being biased into making comparisons by the nature of the pre-test instruction. Cock et al. (1994) raised the possibility that the instruction to rely on memory cues subjects to engage in a post hoc comparison. To test this claim they conducted a standard 3s task but replaced the memory instruction with a “rule” instruction (Cock et al., 1994, Experiment 3). Prior to the test phase of the experiment, subjects were informed that all learning items conformed to a simple rule, and that test items conforming to this rule should be selected. Patterns of performance were similar to those of normal “memory” instruction subjects. In Experiment 2 we control for the possibility of instructional bias on pair type selection. We directly compare the performance of a group given the standard memory instruction at test, with one given an instruction to classify items on the basis of a rule.

Method

Participants. Thirty-two undergraduate students from the University of New South Wales participated in the experiment as part of a third-year practical class. All subjects were aged between 18 and 30 years. All were naïve with respect to the invariant learning paradigm.

Materials. Four sets of learning and test items were created. Sets 1 and 2 used the “late” invariant boundary (times between 6 and 12) and sets 3 and 4 used the “early” invariant (times between 12 and 6). The range of times in each set spanned the entire allowable interval. The basic design of the clocks were similar to those used by Bright and Burton (1994, 1998) though some new designs were added in order to make the induction task more plausible. All clocks featured either Roman or Arabic numerals. Learning items consisted of 30 analogue clock faces printed on four sheets of A4 paper. Test items were pairs of either analogue or digital representations of time. One item in each pair was a novel positive (within the invariant boundary) and one was a negative (outside the boundary). The times in each pair were separated by 1 hour, 3 hours, or 6 hours. Each test set consisted of 24 pairs in total—8 of each of the test pair types (4 digital and 4 analogue). The left/right position of the invariant and the order of presentation of each test pair type were counterbalanced. Analogue and digital pairs were blocked such that subjects either saw 12 analogue pairs followed by 12 digital pairs or vice versa.

Design and procedure. A mixed design was used. There were two within-subjects factors: separation with three levels (1 hour, 3 hour, 6 hour) and test format with two levels (same, changed). The between-subjects factor was instruction type with two levels (memory, rule). The dependent variable was the number of positive items selected at test.

Subjects were run together as part of a practical class. They were presented with the learning items (sets 1–4) equally distributed amongst the class and were told that the study was about clock aesthetics. They were asked to look at each clock in turn and rate it on a five-point scale on how good a representation of time they thought it depicted. Subjects typically took about 5 minutes to complete this task. On completion of the task learning items were removed and subjects were given instructions for the test phase. Those in the “memory” group were told that their memory for the clocks they had rated was to be tested. They were told (falsely) that of the pairs to be presented they had seen one and only one of the times in each pair before. They were asked to choose the time they thought they had seen before and to guess if they were unsure. Subjects in the rule condition were told that all the clocks they had just rated conformed to a simple rule and that one in each to-be-presented pair conformed to that rule and

one did not. They were told to choose the time that conformed to the rule and to guess if they were unsure.

Following the test phase, all subjects were given a post-test questionnaire. The questionnaire from Experiment 1 was modified to make it relevant to the clock stimuli (see the Appendix). Once subjects had written answers to all the questions they were fully debriefed.

Results

Two subjects were removed from the analysis as they revealed explicit knowledge pertaining to the invariant and evidence of the use of this knowledge at test (see below). The mean number of positives selected for the remaining subjects ($N = 30$) was 13.6 (out of a maximum 24), $SD = 3.0$. A single sample t -test showed that this was significantly different from chance performance of 12, [$t(29) = 2.88, p < .05$].

A 2 (Instruction Type: rule, memory) \times 2 (Test Format: same, changed) \times 3 (Separation: 1 hour, 3 hour, 6 hour) mixed model analysis of variance revealed a main effect of separation, [$F(2, 56) = 4.07, p < .05$], but no main effect of test format, [$F(1, 28) = 1.73, p > .05$], or instruction type, [$F(1, 28) = .06, p > .05$]. No interactions reached significance. Mean selection of positives collapsed across the instruction type variable are shown in Table 2. Subsequent analyses were performed on the data collapsed across both the instruction type and test format variables. Planned contrasts revealed a significant difference between the 1-hour and 6-hour pairs, [$F(1, 28) = 8.53, MSe = 1.76, p < .05, (Fc = 6.46)$]. Contrasts between 6- and 3-, and 1- and 3-hour pairs were not significant. Positive selection in the three pair types was compared to chance performance (four in each condition). Selection was significantly different from chance in the 6-hour pairs, [$t(29) = 4.07, p < .001$]. In the 3-hour and 1-hour pairs it was not: 3 hour [$t(29) = 1.08, p > .10$]; 1 hour [$t(29) = 0.48, p > .10$].

Test of explicit knowledge. Out of 32 subjects tested 7 (3 from the rule group and 4 from the memory group) were able to identify accurately the invariant range of times but out of this 7 only 2 indicated using the knowledge of the range at test. This was reflected in their scores—24 out of a possible 24

TABLE 2
Mean selection of positives (out of four) and standard deviations by condition (Experiment 2)

<i>Test format</i>	<i>1 hour</i>	<i>3 hour</i>	<i>6 hour</i>
Same (analogue)	2.03 (1.25)	1.83 (1.21)	2.63 (1.16)
Changed (digital)	2.10 (0.92)	2.50 (0.90)	2.53 (0.90)

selections of the positive. These two subjects were excluded from the main analysis. The mean positive selection of the remaining five subjects was 15.6 (compared to the overall mean of 13.6) suggesting that they were using knowledge of the range even though they claimed to be guessing when asked how they were making their decisions (question 4). In terms of strategies used, some simply reported guessing or relying on instinct and “gut feeling”, whereas others mentioned using familiarity or memory. In total 69% of subjects in the memory group reported using familiarity, compared with only 31% in the rule group. No subject reported the correct rule in response to the question that asked directly what the rule might be.

Discussion

Experiment 2 has demonstrated that the more similar test times are (in terms of hours) the fewer the positives selected at test. We suggested that if subjects were acquiring the 6 to 12 rule, then provided times fell on either side of the invariant boundary selection patterns should be unaffected by the similarity of test pairs. The failure to find this pattern of results demonstrates that subjects have not learnt the experimenters’ rule pertaining to the “6 to 12” time range.

This finding indicates that it is not only the invariant digit task that is open to reinterpretation; performance in the clock task, too, can be accounted for by the learning of correlated information. However, as with the digit task, the nature of this information is not immediately apparent. It is plausible that subjects are learning a less well-defined or fuzzy rule (e.g., “Select later times”) that is only sufficient to produce above chance performance when test times are separated by a 6-hour margin. Equally, subjects could be engaging in a similarity matching process that can operate when test times are dissimilar but breaks down when the similarity between times is increased. We will return to these explanations in more detail in the General Discussion.

The nature of instructions given to subjects before the test phase did not affect performance. This is consistent with Cock et al. (1994, Experiment 3) who found, using the digit task, that selection was unaffected by the instruction to classify on the basis of a rule. Further consistencies between this experiment and Cock et al. (1994) were found in the analysis of verbal reports. In both experiments a lower percentage of subjects instructed to use a rule mentioned relying on familiarity when making test decisions. This indicates that though performance is unaffected by instructions, verbal reports do seem to be influenced by them.

The extended questionnaire revealed a degree of relevant explicit knowledge. However, it was difficult to distinguish between knowledge that was explicit and used at test, from knowledge cued by the questionnaire. It is plausible that subjects make their selections at test believing that they are guessing, and then when prompted to think about the range of times seen on the learning items, are able to come up with the correct answer.

These problems in interpreting data from post-test questionnaires are echoed in many reviews of the implicit learning literature (e.g., Shanks & St. John, 1994) in which questionnaires are criticised for being too insensitive a measure of task relevant knowledge. Although, as Wright and Burton (1995) argue, the relative simplicity of the materials, and the shortness of the test phase in the invariant task, makes it more suited than most to these “insensitive” tests of explicit knowledge, it is still difficult to assess subjects’ awareness of the knowledge guiding their performance *during* the test phase. For this reason, in Experiment 3 we adopt a method for measuring explicit knowledge reported by Dienes, Altmann, Kwan, and Goode (1995) in the artificial grammar learning literature. Instead of inviting subjects to reflect on the basis of their decisions *after* completing the test, we use confidence ratings to record a measure of participants’ subjective confidence in each decision *during* the test phase. This on-line measure circumvents the problem of cuing explicit knowledge with post-test questions, thus allowing us to draw a firmer conclusion about the implicit or explicit status of the acquired knowledge.

Dienes (Dienes et al., 1995; Dienes & Berry, 1997) has suggested that one way in which knowledge may be implicit is in the sense that subjects do not know that they have it. In other words, they lack *metaknowledge* about their knowledge (Dienes & Altmann, 1997; Dienes et al., 1995; Shanks & Johnstone, 1998). If subjects in the invariant task do lack metaknowledge then we can make two predictions. First, subjects should be no more confident in correct decisions than in incorrect decisions (Dienes et al., 1995). Second, subjects should still demonstrate above-chance performance even when they claim to be literally guessing (Cheesman & Merikle, 1986; Dienes et al., 1995). On the other hand if, performance is mediated by potentially explicit knowledge we expect the opposite pattern of results: higher confidence in correct than incorrect decisions, and guess responses associated with chance performance.

EXPERIMENT 3

Experiment 3 followed the same procedure as the memory group from Experiment 2. However, at test, subjects were required to rate their confidence in each selection decision on a five-point scale. Performance of subjects given the usual incidental instructions prior to learning was compared with that of subjects instructed to try and work out the invariant time range. This second group was included so that we could be more confident in drawing conclusions about the knowledge used by subjects in the task (Berry & Cock, 1998).

Method

Participants. Twenty-eight undergraduate students from the University of New South Wales participated in the experiment in return for course credit. All subjects were aged between 18 and 30 years. None had taken part in previous invariant learning experiments.

Materials. In Experiment 2 no differences were found between “early” and “late” invariant times so only late (6 to 12) times were used in Experiment 2. In addition, no differences were found between analogue and digital format test pairs so in this experiment only digital format test pairs were used. Two new “6 to 12” invariant sets were created with corresponding test sets. Test sets consisted of 12 pairs—4 each of the 1-, 3- and 6-hour types. Learning items were printed on A4 paper with a rating scale under each clock. To aid with data collection, test pairs were presented on a computer monitor. Confidence rating scales were printed on a separate piece of paper—1 indicated “complete guess” and 5 “complete certainty”.

Design and procedure. This was a single-factor within-subjects design. The factor was separation with three levels: 1 hour, 3 hour, and 6 hour. The dependent variable was number of positives selected at test. Subjects were given the learning items and asked to rate each clock on how good a representation of time it showed by marking the appropriate phrase on the scale. On completion of this task subjects sat in front of a monitor and were told to choose the time they thought they had seen previously from the to-be-presented pairs. Subjects sat approximately 40 cm from the screen, they made selections by pressing labelled “L” and “R” keys that corresponded to the time on the left or right of the pair. After making each response subjects were instructed to circle a number from 1 to 5 on the confidence-rating scale. In addition to the 28 subjects given the standard incidental instruction, a further 12 subjects (taken from the same subject pool) were given the following explicit instruction prior to the learning phase:

Please look carefully at all the clocks on the sheets of paper in front of you. All the times shown on the clocks fall within a particular range of hours (i.e., between X:00 and Y:00). Please try to work out the range of hours shown on the clocks as this information will help you in a subsequent test.

Once subjects had completed the test phase, both groups were given the same questionnaire as used in Experiment 2 with the addition of one question. The added question read as follows:

- (7) Did you use your knowledge of the time range to help you make your decisions during the test or did you just think of it when answering question 6?

This was included in an attempt to distinguish between knowledge used at test and knowledge cued by post-test questions.

Results and discussion

Subjects in the group that did not receive instructions about the invariant range prior to learning are termed “incidental”, those in the group that did are termed “explicit”. The mean number of positives selected for the incidental group was

7.64 (out of a maximum of 12), $SD = 1.64$. A single sample t -test showed that this was significantly different from chance, [$t(27) = 5.31, p < .001$]. For the explicit group 11.83, $SD = .39$, positives were selected. This was also significantly different from chance, [$t(11) = 51.91, p < .001$]. A single factor repeated measures analysis of variance on the incidental group scores revealed no effect of separation, [$F(2, 54) = 1.60, p < .1$]. Similarly, no effect was found for the explicit group. Mean positive selection in the three pair types for the incidental group is shown in Table 3. Positive selection was compared to chance performance (two in each condition). This revealed that selection was above chance for the 6-hour pairs, [$t(27) = 3.73, p < .001$], and the 3-hour pairs, [$t(27) = 3.73, p < .001$], but not for the 1-hour pairs, [$t(27) = 1.61, p > .1$]. For the explicit group, selection was above chance for all three pair types.

Confidence ratings. The minimum confidence rating was 1 and the maximum was 5. Overall mean confidence for the incidental group was 2.32, $SD = .78$; for the explicit group it was 4.10, $SD = .58$. Mean confidence for the three pair types for the incidental group is shown in Table 4. A 1×3 Separation (1 hour, 3 hour, 6 hour) within-subjects ANOVA on the incidental group confidence data revealed a main effect of separation, [$F(2, 54) = 6.33, p < .05$]. Planned contrasts revealed significant differences between confidence on the 1- and 3-hour pairs, [$F(1, 27) = 9.19, MSe = .24, p < .05$]. Contrasts between the 1- and 6-hour pairs were not significant.

Confidence in incorrect and correct decisions was averaged for all incidental subjects. A difference in confidence between correct and incorrect decisions was taken as a measure of metaknowledge. Mean confidence in correct decisions was 2.43, $SD = .76$ and in incorrect it was 2.17, $SD = .86$. A paired samples t -test

TABLE 3
Mean selection of positives (out of four) and standard deviations (Experiment 3)

<i>Test format</i>	<i>1 hour</i>	<i>3 hour</i>	<i>6 hour</i>
Changed (digital)	2.29 (0.90)	2.64 (0.90)	2.71 (1.00)

TABLE 4
Mean confidence rating (out of five) and standard deviation for three test pair conditions (Experiment 3) (incidental group)

<i>Separation</i>	<i>1 hour</i>	<i>3 hour</i>	<i>6 hour</i>
	2.10 (0.87)	2.50 (0.88)	2.32 (.79)

showed a significant difference between these means, [$t(26)^1 = 2.13, p < .05$]. In the explicit group only three out of twelve subjects made incorrect decisions and these subjects only made one incorrect decision each. For this reason a comparison of correct and incorrect decisions was not appropriate. Mean level of confidence in correct decisions was 4.09, $SD = .57$ and in incorrect decisions it was 3.67, $SD = 2.31$.

For the incidental group, responses given a confidence rating of 1 (complete guess) were analysed to see if performance differed from chance. If C represents the number of Correct responses thought to be a guess and I the number of Incorrect responses thought to be a guess, then C-I gives us a measure of low confidence knowledge (Dienes et al., 1995). A score of 0 would indicate random responding. Mean value of C-I was .91, $SD = 2.67$; a single sample t -test revealed that this was not significantly different from 0, [$t(21)^2 = 1.60, p > .1$]. The upper limit of the 95% confidence interval for this difference was 2.09.

Test of explicit knowledge. One subject from the incidental group correctly identified the range of times, but in response to question 7 said that they had only thought about it after reading question 6. This subject's score was 7 and their overall confidence was 2.75. Out of 28 subjects, 19 mentioned using familiarity or trying to visualise learning items. Many subjects expressed surprise when told about the invariant time range during debriefing. In contrast, all 12 subjects in the explicit group correctly identified the range of times and said that they had used knowledge of the range to make their decisions at test.

The aim of Experiment 3 was to assess the implicit or explicit status of the knowledge underlying performance in the clock task. Confidence rating data showed that incidental subjects were more confident in their correct decisions than in their incorrect decisions. This suggests that subjects were aware of when they were applying knowledge as opposed to merely guessing (Dienes & Altmann, 1997). In other words, subjects possessed metaknowledge. Furthermore, for the decisions in which subjects claimed to be literally guessing (a confidence rating of 1) responses were no different from random responding. This random responding implies that there is no underlying influence of implicit knowledge. In order to determine the robustness of this interpretation of the guessing criterion, we examined the sensitivity of the test to pick up the influence of implicit knowledge. The mean number of guess responses for subjects in the incidental group was 3.85 (out of a possible 12). The mean invariant selection for this group was 7.64/12, thus the maximum expected number of correct responses is $(7.64/12) \times 3.85$, and the maximum expected number of incorrect responses is $(4.36/12) \times 3.85$. Given this, the maximum mean expected difference score is $(7.64 - 4.36) \times 3.85/12 = 3.28 \times 0.32 = 1.05$. This analysis

¹There was missing data for one subject who made no incorrect decisions.

²Missing data for six subjects who gave no confidence ratings of 1.

demonstrates that the maximum mean expected difference score (1.05) lies inside the upper limit of the 95% confidence interval (2.09) reported earlier.³ We are therefore unable, on the basis of the guessing criterion, to rule out the presence of implicit knowledge. Furthermore, we acknowledge that possible individual differences in the interpretation of the scale and in willingness to report low confidence knowledge may have resulted in our measure being a rather conservative estimate of implicit knowledge.

Post-test questionnaires, which revealed minimal verbal knowledge of the invariant, suggest the presence of some implicit knowledge. It seems that, although some decisions were associated with higher levels of confidence, the underlying knowledge was not fully available to verbal report and thus may have been implicit. Explicit subjects, who were provided with a strategy for success, performed, as expected, better than incidental subjects (in terms of positive selection) and exhibited higher levels of confidence and relevant explicit knowledge.

The failure to find a main effect of separation in the positive selection data is inconsistent with Experiment 2. This failure is probably due to a loss of power resulting from the reduction in the number of test pairs from 24 to 12. Support for this conclusion comes from the finding of a significant main effect of separation across Experiments 2 and 3, [$F(2, 112) = 3.91, p < .05$], and a non-significant Separation (1, 3, 6 hour) \times Experiment (Experiment 2, Experiment 3) interaction, [$F(2, 112) = .111, p > .1$]. Mean trends and comparisons to chance are also consistent with this explanation.

Also somewhat counter-intuitive is the finding that subjects were most confident in decisions concerning the 3-hour separation pairs. One possible explanation for this is that some 6-hour pairs contained items in which both times were close to the invariant boundary (e.g., 6:50 vs 12:50), this may have confused some subjects leading to lower levels of confidence. In summary, Experiment 3 has provided stronger evidence that performance in the invariant task is not mediated solely by implicit knowledge.

GENERAL DISCUSSION

On first inspection, the invariant learning paradigm seems to provide straightforward evidence for the implicit learning of an abstract rule. Reanalysis of the original phenomenon has resulted in a reconsideration of that conclusion. There is now a growing body of evidence suggesting that performance in the invariant learning tasks does not rely on subjects' implicitly acquiring the rules that it was originally assumed were learnt (Churchill & Gilmore, 1998; Cock et al., 1994; Wright & Burton, 1995). The three experiments presented here contribute to that body of evidence. We argue that our findings are difficult to reconcile under the

³We thank Referee 2 for suggesting this analysis.

original explanations of the invariant digit and clock task. In Experiment 1, if subjects performed purely on the basis of the invariant 3 rule then the manipulation of repeated digits would have had no effect. Similarly, in Experiment 2, performance based on a “6 to 12” rule predicts no difference between the selection of 6-hour and 1-hour pairs. Finally, processing mediated by implicit knowledge predicts no relationship between confidence and accuracy in decisions. In all these instances, we found significant effects.

The representation of knowledge

The finding that subjects are not learning the original experimenters’ rule leaves the issue of what knowledge subjects are acquiring open to question. The results of Experiment 1 suggest, consistent with the findings of Wright and Burton (1995), that subjects learn about the relative distribution of repetitions in learning and test strings and that this knowledge enables subjects to reject the distinctive repetition containing strings. Wright and Burton tentatively concluded that this knowledge was instantiated as an explicit rule. We suggest that the pattern of rejection equally may be due to subjects making judgements based on an analogy to learning items. It is plausible that over the course of the test phase subjects abstract the “reject repetitions” rule, but in order for this abstraction to occur, the initial stages of the test phase (e.g., the first three or four pairs) must involve a process of making similarity matches with memories of learning strings. A subject engaging in such a process will tend to reject the items containing repetitions, as they are highly distinctive and thus less similar to the majority of learning items that do not contain repetitions (Wright & Burton, 1995).

We argue that this more general form of similarity matching also provides a good explanation of performance in the clocks task. As suggested by Cock et al. (1994) it is plausible that because of possible ambiguities in the placing of hands on the clock faces, subjects may encode times seen in the learning phase in a rather approximate fashion. Thus, when confronted with a 1-hour pair (e.g., 5:50 vs 6:50) at test, comparison with a partial memory for a learning time “around 5 or 6” would result in the inability to discriminate the positive (6:50) from the negative (5:50) time. In contrast, a subject presented with a 6-hour pair (e.g., 8:25 vs 2:25) may have a partial memory for a time “around 8:30” and in addition have the feeling that times around 2 were *not* seen. Thus, subjects may either accept the positive time actively, on the basis of its similarity to partial memories of learning items, or passively through the rejection of an item that is dissimilar to any previously encountered times. The removal of the dissimilar time in the 1-hour pairs renders this comparison process ineffective and thus results in the inability to perform the task successfully. Subjects’ verbal reports that include statements of the form “I know I did not see time X” support this explanation (see Churchill & Gilmore, 1998, for a similar argument).

In summary, we are not able to distinguish equivocally between rule-based and similarity-based processing but, at the very least, we argue that the initial stages of the test phase must involve some similarity matching. However, it is possible that over the course of the test phase a “reject repetitions” or a fuzzy “select later times” rule may be abstracted. It is perhaps, therefore, safest to conclude that the invariant tasks are most likely to involve a combination of rule-based and similarity-based processing. This conclusion is consistent with computational models of other implicit learning paradigms, e.g., sequence learning (Cleeremans, 1994), findings from artificial grammar learning tasks (Vokey & Brooks, 1994), and studies of category learning (Nosofsky, Clark, & Shin, 1989) that all provide evidence for both rule and similarity-based processing.

Finally, the cross-format transfer effects found in the three experiments indicate that the underlying representation of knowledge includes more than simple surface feature information. However, in itself, this finding does not enable us to distinguish between rule- and similarity-based processing, because evidence for abstract representations is equally consistent with both forms of processing (Hahn & Chater, 1998).

Implicit and explicit knowledge

We argued that the finding that subjects were more confident on trials in which they were correct than on those in which they were incorrect indicated the presence of some metaknowledge. Furthermore, analysis of subjects’ guessing responses failed to reveal any influence of implicit knowledge. To conclude from these findings that invariant learning is based exclusively on explicit knowledge would be over-interpreting our data. The discrepancy between subjects’ verbalisable knowledge and their confidence, the overall low confidence of incidental subjects (in comparison to explicit subjects), and their genuine surprise during debriefing, all suggest otherwise. We acknowledge that we may have been unable to detect implicit knowledge due to a lack of power and because of the difficulties inherent in dissociating implicit from explicit knowledge. Invariant learning, as with many of the tasks used to investigate implicit learning, suffers from the problem that the response focused upon (selection of the positive) is not neutral with respect to potential contamination by explicit knowledge (Vinter & Perruchet, 1999). Performance based on implicit knowledge results in the same outcome—above-chance selection of the positive—as performance based on explicit knowledge of the invariant characteristic (or correlated information). Consequently, dissociating the influence of implicit and explicit knowledge on performance is problematic (see Jacoby, 1991; Jacoby, Toth, & Yonelinas, 1993). This difficult issue is beginning to be addressed in the literature through the development of new experimental tasks. For example, Vinter and Perruchet (1999) reported an innovative drawing task in which they were able to demonstrate the influence of implicit processes because the observed adaptations

in drawing practice could not be expected by conscious exploitation of explicit knowledge. Though our results do not deny the existence of implicit knowledge, we can conclude (from our finding of metaknowledge) that, contrary to earlier interpretations (e.g., McGeorge & Burton, 1990; Bright & Burton, 1994), invariant learning is not mediated purely by implicit knowledge.

In summary, we believe that our data are most consistent with the view that subjects acquire a mixture of implicit and explicit knowledge (see Dienes & Berry, 1997). This view is echoed by the findings of Dienes et al. (1995), who report that in an artificial grammar learning task subjects in three experiments exhibited both implicit knowledge, evidenced by above-chance performance when claiming to be guessing, and explicit knowledge through greater confidence in correct decisions. Shanks and Johnstone (1998, 1999) report similar findings using the sequence learning task, though in sequence learning the evidence for implicit knowledge appears to be much weaker.

Conclusions

The experiments reported here reflect a shift that can be seen in many recent studies of implicit learning. As noted by Cleeremans, Destrebecqz, and Boyer (1998) in a review, early claims for a sophisticated unconscious mechanism capable of abstracting and applying rules have been replaced by accounts emphasising the importance of task demands during learning and the congruence between learning and test conditions (Whittlesea & Wright, 1997; Wright & Whittlesea, 1998). In accord with this shift in perspective, we have presented further stronger evidence against the implicit abstraction of the experimenters' rules in the invariant tasks. We propose, consistent with the learning underlying many so-called implicit learning tasks, that the invariant tasks involve a mixture of both implicit and explicit knowledge and a combination of similarity- and rule-based processing. Future investigations need to develop more powerful tasks, which will allow us to assess the relative contributions of these types of knowledge and forms of processing to implicit learning.

Manuscript received August 1999

Revised manuscript received September 2000

REFERENCES

- Berry, D., & Cock, J. (1998). Implicit learning of invariant features? In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp. 135–159). Thousand Oaks, CA: Sage.
- Blaxton, T.A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 657–668.
- Bright, J.E.H., & Burton, A.M. (1994). Past midnight: Semantic processing in an implicit learning task. *Quarterly Journal of Experimental Psychology*, *47A*, 71–89.

- Bright, J.E.H., & Burton, A.M. (1998). Ringing the changes: Where abstraction occurs in implicit learning. *European Journal of Cognitive Psychology*, *10*, 113–130.
- Brooks, L.R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorisation*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Cheesman, J., & Merikle, P.M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, *40*, 343–367.
- Churchill, E.F., & Gilmore, D.J. (1998). Selection through rejection: Reconsidering the invariant learning paradigm. *Quarterly Journal of Experimental Psychology*, *51A*, 1–17.
- Cleeremans, A. (1994). The representation of structure in sequence prediction tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and unconscious information processing* (pp. 783–809). Cambridge, MA: MIT Press.
- Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*, 406–416.
- Cock, J.J., Berry, D.C., & Gaffan, E.A. (1994). New strings for old: The role of similarity processing in an incidental learning task. *Quarterly Journal of Experimental Psychology*, *47A*, 1015–1034.
- Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains: How implicit and how abstract? In D.C. Berry (Ed.), *How implicit is implicit learning?* (pp. 107–123). Oxford, UK: Oxford University Press.
- Dienes, Z., Altmann, G.T.M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1322–1338.
- Dienes, Z., & Berry, D. (1997). Implicit learning: Below the subjective threshold. *Psychonomic Bulletin and Review*, *4*, 3–23.
- Gomez, R.L. (1997). Transfer and complexity in artificial grammar learning. *Cognitive Psychology*, *33*, 154–207.
- Gomez, R.L., & Schvaneveldt, R.W. (1994). What is learned from artificial grammars? Transfer tests of simple association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 396–410.
- Hahn, U., & Chater, N. (1998). Similarity and rules: Distinct? exhaustive? empirically distinguishable? *Cognition*, *65*, 197–230.
- Jacoby, L.L. (1991). A process dissociation framework: Separating unconscious from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Jacoby, L.L., & Hayman, C.A. (1987). Specific visual transfer in word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 456–463.
- Jacoby, L.L., Toth, J.P., & Yonelinas, A.P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, *122*, 139–154.
- McGeorge, P., & Burton, A.M. (1990). Semantic processing in an incidental learning task. *Quarterly Journal of Experimental Psychology*, *42A*, 597–609.
- Nosofsky, R., Clark, S., & Shin, H. (1989). Rules and exemplars in categorization, identification and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 282–304.
- Perruchet, P., & Gallego, J. (1997). A subjective unit formation account of implicit learning. In D.C. Berry (Ed.), *How implicit is implicit learning?* (pp. 124–161). Oxford, UK: Oxford University Press.
- Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *19*, 264–275.
- Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *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.
- Shanks, D.R., & Johnstone, T. (1998). Implicit knowledge in sequential learning tasks. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of implicit learning* (pp. 533–572). Thousand Oaks, CA: Sage.

- Shanks, D.R., & Johnstone, T. (1999). Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1435–1451.
- Shanks, D.R., & St. John, M.F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences*, *17*, 367–447.
- Vinter, A., & Perruchet, P. (1999). Isolating unconscious influences: The neutral parameter procedure. *Quarterly Journal of Experimental Psychology*, *52A*, 857–875.
- Vokey, J.R., & Brooks, L.R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328–344.
- Vokey, J.R., & Brooks, L.R. (1994). Fragmentary knowledge and the processing-specific control of structural sensitivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1504–1510.
- Whittlesea, B.W.A., & Wright, R.L. (1997). Implicit (and explicit) learning: Acting adaptively without knowing the consequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 181–200.
- Wright, R.L., & Burton, A.M. (1995). Implicit learning of an invariant: Just say no. *Quarterly Journal of Experimental Psychology*, *48A*, 783–796.
- Wright, R.L., & Whittlesea, B.W.A. (1998). Implicit learning of complex structures: Active adaptation and selective processing in acquisition and application. *Memory and Cognition*, *26*, 402–420.

APPENDIX

Post-test questionnaire used in Experiment 1

- (1) Were any of your responses guesses?
- (2) Were you aware of using any strategies when guessing? What were the strategies?
- (3) Did you notice anything systematic about the numbers in the first part of the experiment?
- (4) Did any of the numbers appear more frequently than any others?
- (5)* One number appeared more frequently than any other during the first part of the experiment. Could you circle the number you think it was: If you're not sure please guess.

1 2 3 4 5 6 7 8 9

Post-test questionnaire used in Experiments 2 and 3

- (1) What do you think the experiment was about?
- (2) Did you notice any peculiarities or common features about the times shown on the clocks that you rated initially?
- (3) Did you notice any peculiarities or common features about the times presented as pairs in the second half of the experiment?
- (4) How did you make your decisions when choosing one time from each pair?
- (5) All the times shown on the clocks that you rated initially conformed to a simple rule. Have you any idea what that rule is?
- (6)* All the times shown on the clocks that you rated initially fell within a particular time range. Have you any idea what that range is?

* These questions were printed on the reverse side of the sheet to prevent cueing of knowledge about the nature of the invariant.