cations: The digit pairs 76, 65, 54, and 43 each occurred in two items. The other two digits were selected so that the combination had not been shown in training, no item contained repeated digits, and their sum was the same as the sum of digits in those locations in at least one training item (e.g., 7265). The other member of each pair violated the descent-by-one pattern; otherwise, its digits were selected subject to the same requirements as the first set.

Procedure. Half of the subjects were asked to add the first pair of digits in each training stimulus, as well as the last pair; the other half added the inside pair of digits, and also the outside pair. At test, all subjects were (falsely) told that one member of each test pair had been seen in training and were asked to identify the old member of each pair.

Results and Discussion

Overall, subjects claimed to recognize target items (containing the invariant descending-by-one pattern) on 49% of trials, against the chance rate of 50% [$t(19) < 1$]. The specific induction task made little difference. Subjects required to add the last two digits together selected target items on only 45% of trials, and subjects who had processed those digits in separate addition operations selected target items on 53% of trials; neither was reliably different from chance [$t(9) < 1$ in both cases], nor were they reliably different from each other [$t(18) = 1.2$]. We concluded that there was no evidence that subjects had acquired any ability, explicitly or implicitly, to discriminate items preserving the invariant pattern from those violating it.

This failure demonstrates that sensitivity to an abstract invariant is not an automatic consequence of processing stimuli extensively. Instead, that sensitivity appears to depend on the relationship between the invariant and the knowledge the subject computes under the control of the induction task. In Experiment 2A, the information that subjects were required to compute (the sums of digit pairs) was directly related to the abstract invariant (the equality of sums of pairs); although not asked to compare the sums of the digit pairs, those subjects became explicitly aware of the invariant relationship. In Experiment 2B, the information that subjects were required to compute (again the sums of digit pairs) was unrelated to the invariant (the constant difference of the terminal digits, or their even–odd relationship); those subjects did not become sensitive to the invariant. We suspect that these are examples of a general principle—that people only learn about those properties of their experience that they are led to process to satisfy the task or that are systematically related to their task, and do not automatically and unconsciously compute abstract properties of the domain that are irrelevant to their task.[3]

How, then, do people become sensitive to abstract stimulus properties without becoming aware that they exist? We suggest that people do not directly learn about those properties. Instead, in the course of satisfying the induction demand, they learn some information that is correlated with the implicit property. In that case, the subjects could discriminate test items preserving an implicit invariant from those violating the invariant, without ever realizing that the invariant exists, but they would do so using a different form of information. That is, the subjects become accidentally sensitive to the invariant, through its correlation with properties that they have directly processed under the control of the induction task. We illustrated this principle in the next experiment, showing that subjects can become sensitive to an abstract invariant without awareness, but only if they are directly required to process some information that is indirectly related to that invariant.

Experiment 3
Implicit Sensitivity as an Incidental By-Product of Explicit Learning

Experiment 3 was designed to demonstrate a case of truly implicit learning in which subjects actually become sensitive to an abstract invariant without becoming aware of its existence. However, the experiment also shows that this sensitivity does not occur through automatic abstraction of the invariant, occurring independently of the demands of the induction task. Instead, that sensitivity is a by-product of computing information to satisfy those demands. Specifically, this sensitivity demonstrates that (1) different induction tasks cause subjects to compute different types of information about the stimuli; (2) subjects only compute information of direct relevance to their task, and do not unconsciously compute information about structural invariants that are not relevant; but (3) the information that subjects compute directly, in response to the demands of the task, can make them indirectly sensitive to such invariants, producing the phenomenon of "implicit learning."

To illustrate these points, we created a set of 16 four-digit numbers, each of which followed the pattern odd–even–odd–even (e.g., 1834). This invariant odd–even pattern was the implicit rule of the set. Subjects were never told about the existence of this invariant, either in training or in test, and the induction tasks did not require subjects to notice or compute any information about evenness or oddness. The question was whether subjects exposed to these stimuli would become sensitive to that invariant without becoming aware of it, and if so, whether that sensitivity occurred through an autonomous, unconscious abstraction process or, alternatively, because the induction task led the subjects to compute information that was indirectly related to the invariant.

To discover whether the subjects' sensitivity to an invariant depends on the specific nature of the induction task, we varied that task between groups. One group was asked to read each number aloud, pronouncing each as a pair of two-digit numbers (e.g., to read 1258 as "twelve fifty-eight"); the other was to read each number as four separate digits (e.g., "one–two–five–eight"). Both of these tasks expose the subject to the entire structure of each stimulus; thus, if implicit learning consists of automatic abstraction of general properties of the domain, we would expect subjects to become as sensitive to the odd–even rule in one condition as in the other. However, if the knowledge that supports sensitivity to implicit properties of the domain actually consists of information directly and explicitly computed to satisfy the demands of the induction task, then we would expect differential sensitivity to the implicit invariant after the two types of training.

The latter prediction is predicated on the idea that the two induction tasks actually caused subjects to encode the stimuli in different ways. We checked that in Experiment 3A. We then tested the major hypothesis of the study, that differences in the induction task would alter the subjects' sensitivity to the implicit invariant, in Experiment 3B, presented later.

Experiment 3A

Method

Subjects. Twenty undergraduate students attending Simon Fraser University participated in Experiment 3A, for course credit.

Materials. To construct stimuli for the induction phase, we created a pool of bigrams, each consisting of an odd digit followed by an even digit, using only the digits from 1 to 8. The set of bigrams used consisted of {12, 16, 34, 38, 52, 58, 74, 76}. Each bigram was combined with four others to create four-digit stimuli. Each bigram occurred twice in the first half of a stimulus (e.g., 1258 and 1234) and twice in the second half (e.g., 3812 and 7412). No digit was allowed to occur twice in any stimulus. This process generated 16 stimuli. Across these stimuli, each single digit occurred at two stimulus locations (depending on whether it was even or odd), and with the same frequency as all other digits.

Two types of stimuli were generated for test. The first type was created by reversing the bigrams of half the training items; for example, 1258 produced 5812, and 7412 produced 1274. Each bigram from the training was used twice, once in the first half and once in the second half of a test item. This produced eight stimuli, consisting of the same bigrams used to create training stimuli, but presenting those bigrams in a novel combination.

The second set of stimuli was created using a set of bigrams that had not been used earlier, namely {14, 18, 32, 36, 54, 56, 72, 78}. Each was used twice, once in the first half and once in the second half of a test item, again producing eight stimuli. Although this set consisted of bigrams that were novel at test, they consisted of the same set of single digits (from 1 to 8), presented at the same locations and with the same frequency, as in the previous set. The only difference between stimulus types was the familiarity of their component bigrams from the training phase.

Procedure. In a training phase, subjects were shown the 16 training stimuli, one at a time, in random sequence, and asked to memorize each for a later test. One group was instructed to perform the memorization by saying each number aloud, reading it as two two-digit numbers (e.g., 1258 read as "twelve, fifty-eight"). The other group was instructed to read each number aloud as a string of separate digits. In addition, the other group was asked to judge each number as low or high, depending on whether it was 5 or larger (e.g., 1258 read as "1–low–2–low–5–high–8–high"). Both groups were allowed 5 sec to rehearse each item and were given two passes through the stimuli.

At test, subjects were told that they would see some new stimuli and some stimuli repeated from the training and were asked to indicate which they had seen before. The 16 (actually all novel) test items were presented, one at a time, in randomized sequence.

Results and Discussion

The two induction tasks were intended to cause subjects to encode the numbers differently, as digit pairs and as lists of individual digits, respectively. This experiment served as a manipulation check, permitting us to evaluate whether the two types of induction task really had caused subjects to encode the stimuli differently. Subjects who had only encoded information about individual digits (including the location in which those digits occurred) would be unable to discriminate between the two types of test items because the two types presented the same digits in the same locations with equal frequency. In contrast, subjects who had encoded specific pairs of digits presented by training items would be able to discriminate between the two types of test items because one presented those same pairs (recombined to make new items) and the other presented novel pairs. In effect, these subjects could claim to recognize test items by recognizing their parts. Knowledge about the odd–even rule, whether acquired directly or indirectly, was not an issue in this test because all test stimuli followed that rule.

We observed that subjects asked to read training items as single digits could discriminate the two types of test item to some degree. They claimed to recognize 56% of items consisting of old digit pairs and rejected 59% of items containing new pairs. Treating the former as hits and the latter as correct rejections, they achieved an overall 57.5% ability to discriminate the two types of item, reliably above chance [$t(9) = 2.88, p < .018$]. We concluded that this group had acquired some knowledge about bigrams in spite of our attempt to make them process training items as separate digits.[4]

However, we also observed that the other group of subjects, who read training items as two pairs of digits, were better able to discriminate the two types of test items. They accepted 64% of items consisting of old pairs and rejected 78% of items containing new pairs. They thus achieved an overall 71% ability to discriminate the two types of items, reliably greater than the first group [$t(18) = 2.36, p < .030$]. We concluded that this group had learned more extensively about the bigrams of training items than had the first group. We also noted that both groups achieved discrimination between the types more by rejecting items containing new bigrams than by accepting items containing familiar bigrams.

Now that we knew that the induction tasks caused subjects to process and encode different information about the training stimuli, the question was whether this difference would have any effect on their sensitivity to the odd–even rule. According to the "automatic abstraction hypothesis," the different induction tasks should have no effect on discrimination of legal from illegal items: Both tasks should merely serve as opportunities to observe and abstract the invariant properties of the domain. In contrast, if subjects discriminate legal from illegal items by accepting familiar bigrams and rejecting unfamiliar ones, then the more extensive bigram encoding of the bigram-reading group should enable them to discriminate legal from illegal items more effectively. We tested this issue in Experiment 3B.

Experiment 3B

Method

Subjects. Twenty undergraduate students attending Simon Fraser University participated in Experiment 3B, for course credit.

Materials. Training stimuli were identical to those of Experiment 3A. Thus each training item followed the pattern odd–even–odd–even.

We created two sets of test stimuli, one of which was identical to the "familiar bigrams" set used in Experiment 3A. This set necessarily followed the odd–even rule and so qualified as the "legal" set.

A second set of eight "illegal" stimuli was created by the same bigram-combination rules as the first set; like the first set, these items consisted of familiar bigrams, but presented a novel combination of those bigrams. Each item of this set was then modified by the replacement of one digit by another. In each case, an even digit was replaced by an odd digit, or vice versa. In consequence, the item now violated the odd–even rule at one of its four locations (e.g., 1258 modified to 1268, producing a violation of the rule in the third location). Violations occurred with equal frequency at each location across the set.

Procedure. The training and test phases were conducted as in Experiment 3A. Following the test, subjects were interviewed about the basis of their recognition decisions.

Results and Discussion

The group of subjects who read numbers as lists of individual digits demonstrated sensitivity to the implicit rule (the invariant odd–even pattern) in the recognition test: They claimed to recognize 51% of legal stimuli and rejected 65% of illegal stimuli. Their overall discrimination of legal from illegal stimuli was thus 58%, reliably greater than chance [$t(9) = 3.28$, $p < .009$]. Asked how they had performed their recognition judgments, the subjects reported using feelings of familiarity and/or a variety of mini-rules, such as "big numbers at the beginning" or "if it had an eight"; none reported noticing the odd–even pattern that defined legality.

The group of subjects who had read numbers as digit pairs claimed to recognize 60% of legal items and rejected 80% of illegal items. Like the first group, they ascribed their judgments to familiarity and to a host of mini-rules unrelated to the actual implicit odd–even rule. Overall, they achieved 70% discrimination, reliably greater than that of the other group [$t(18) = 3.31$, $p < .004$].

We thus observed, in both groups, the standard "implicit learning" effect, that subjects can become sensitive to a structural invariant without becoming aware of doing so. However, there is evidence that that sensitivity was not achieved through direct, unconscious abstraction of the invariant during the training, but instead through the encoding of bigrams and use of that knowledge to reject test items containing unfamiliar bigrams. First, according to the automatic abstraction hypothesis, implicit learning is supposed to be unselective and stimulus driven (see discussion after Experiment 1). If that were the case, we should observe as much sensitivity to the abstract invariant following the single-digit-encoding task as the bigram-encoding task: Both groups were exposed to the full structure of each item and of the entire domain. Contrary to that hypothesis, the degree of sensitivity to the rule clearly depended on the nature of the induction task: Subjects instructed to read numbers as bigrams achieved considerably greater discrimination between legal and illegal items.

Second, we knew from Experiment 3A that the bigram induction task encouraged more extensive encoding of bigrams than did the single-digit task, permitting subjects greater sensitivity in rejecting items containing unfamiliar bigrams. The data of Experiment 3B demonstrated an almost identical pattern of performance: greater discrimination following the bigram induction task and success based more heavily on rejecting items containing unfamiliar bigrams than accepting items containing familiar bigrams. We concluded that subjects in Experiment 3B performed at test in the same way as subjects in Experiment 3A: They claimed to recognize items on the basis of the familiarity of their bigram components and succeeded to the extent that the induction phase had required them to compute such information. Their apparent sensitivity to the implicit invariant was indirect and accidental, resulting from the fact that violations of the rule occurred only in unfamiliar bigrams.

This result demonstrates our central thesis, that the induction task of an implicit learning experiment is not a neutral opportunity to observe and unconsciously abstract the general structure of a domain, thereby becoming implicitly sensitive to its abstract properties. Instead, it clearly demonstrates that the development of sensitivity to implicit aspects of structure is a direct consequence of the demands of the induction task. It further demonstrates that the ability to discriminate between test items that do and do not violate the implicit rule need not be supported by direct knowledge of that rule, but can instead be granted by knowledge correlated with that rule. Finally, it demonstrates that the effective knowledge that permits the subject to show sensitivity to an implicit rule is not the result of unselective, stimulus-driven and autonomous abstraction, but is instead computed directly in response to the demands of the induction task.

On this understanding of implicit learning, during the induction phase subjects only compute information that is relevant to their current purpose for encountering stimuli. In the test phase, the experimenter varies test stimuli along a dimension that subjects cannot anticipate and about which they have no direct knowledge. The subjects apply whatever knowledge they have acquired, and hope for the best. If the knowledge they computed in training is correlated with the dimension manipulated by the experimenter, then their performance will be above chance in the test; if not, they will be at chance on that discrimination. Finding above-chance performance in the test