Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 5, 846–861

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.5.846

Error-Driven Knowledge Restructuring in Categorization

Michael L. Kalish
University of Louisiana at Lafayette

Stephan Lewandowsky and Melissa Davies
University of Western Australia

Knowledge restructuring occurs when people shift to a new strategy or representation during learning. Although knowledge restructuring can frequently be experimentally encouraged, there are instances in which people resist restructuring and continue to use an expedient but imperfect initial strategy. The authors report 3 category learning experiments that reconciled those conflicting outcomes by postulating that, for restructuring to occur, learners must be dissatisfied with their knowledge and a usable alternative must be available. In line with expectation, restructuring was elicited only when an alternative strategy was pointed out and when people's initial expedient strategy entailed performance error. Neither error nor information about the alternative strategy by itself was sufficient to induce restructuring.

Keywords: category learning, strategy change, task difficulty, selective attention, error-driven learning

The best way to master a complex novel task is to maximize performance as quickly and reliably as possible. To this end, one may initially process stimuli in expedient ways, by focusing on a limited amount of information and ignoring the rest. Although this restricted initial focus can accelerate immediate learning, it often supports only mediocre performance. To excel, learners need to *restructure* their knowledge either by incorporating previously ignored aspects of the stimuli or by developing a better strategy to deal with the task. For example, children might begin doing arithmetic by counting on their fingers, before they switch to a simpler and more efficient memory-retrieval strategy for the same problems after additional practice (Koshmider & Aschcraft, 1991; Shrager & Siegler, 1998). When confronted with an unfamiliar arithmetic task (e.g., A + 3 = ?), adults likewise may first resort to counting (e.g., "B . . . C . . . D") before switching to a memory-based strategy (e.g., Logan, 1988).

The idea that knowledge is restructured during learning has been formalized many times and forms an integral part of several theories of skill acquisition (e.g., Compton & Logan, 1991; Logan, 1988; Rumelhart & Norman, 1978; Shrager & Siegler, 1998). Although the concept of restructuring is allied with diverse theoretical antecedents (e.g., the Piagetian view of human development or the Kuhnian view of the history of science), at least one common attribute applies to all conceptions of restructuring: Knowledge restructuring invariably involves a shift from one strategy or representation to another that is qualitatively and identifiably different, without that shift necessarily being accompanied or triggered by a change in the stimuli. The shift between strategies or representations differentiates restructuring from other forms of learning, which we call *performance improvement*, that involve the mere consolidation of existing knowledge through fine-tuning and the more efficient application of otherwise unchanged processes. For the purposes of this article, we thus define knowledge restructuring as an empirically identifiable shift from one cognitive strategy to another.

Notwithstanding the importance of the concept and its wide applicability, the cognitive processes underlying restructuring are not fully understood (Lewandowsky, Kalish, & Griffiths, 2000). This article examines two potentially important factors, namely, the role of performance error and availability of an alternative strategy. Our investigation focuses on category learning, a paradigm that is particularly suitable for the identification and comparison of different strategies. We proceed as follows. First, we provide a broader context for our work by touching briefly on knowledge restructuring in naturalistic settings such as in the classroom or across developmental stages. This brief survey reveals that knowledge restructuring is often triggered by the learners' dissatisfaction with their current knowledge. We then turn to category learning and examine the limited existing literature. This examination reveals two conflicting outcomes, with restructuring occurring in some situations but not in others. We propose that these conflicting findings can be reconciled by extending the dissatisfaction concept from naturalistic settings to an error-driven learning account of categorization. We then present three category learning experiments in which we encourage knowledge restructuring by revealing additional task-relevant information during learning. As predicted by the error framework, restructuring could not be induced when performance error was low, either because cue validity was high or because crucial exemplars could be memorized. By contrast, when error was present, restructuring occurred if, and only if, additional information about the task was revealed. We conclude that performance error and the availability of a usable alternative are individually necessary and jointly sufficient to induce knowledge restructuring in categorization.

Michael L. Kalish, Institute of Cognitive Science, University of Louisiana at Lafayette; Stephan Lewandowsky and Melissa Davies, School of Psychology, University of Western Australia, Perth, Australia.

Preparation of this article was facilitated by a Discovery Grant from the Australian Research Council to Michael L. Kalish and Stephan Lewandowsky. We thank Leo Roberts for his assistance during data collection and manuscript preparation. We also thank Mathew Duncan for his comments on an earlier version of the article.

Correspondence concerning this article should be addressed to Michael L. Kalish, Institute of Cognitive Science, University of Louisiana at Lafayette, Lafayette, LA 70504-3772. E-mail: kalish@louisiana.edu

Knowledge Restructuring in Natural Settings

Knowledge restructuring underlies many of the conceptual changes that occur throughout the life span. Accordingly, research in naturalistic settings has performed analyses of learning in various domains. For example, Pearsall, Skipper, and Mintzes (1997) looked at the progressive changes in structural complexity of the topical knowledge held by college biology students. Each participant was required to construct a "concept map" of his or her understanding of the subject at 4-week intervals throughout a semester. The maps were then rated for complexity and structural changes. The results showed that, during learning, people abandoned one knowledge structure and assembled another. Similar results have been found in other naturalistic domains, including in primary and high school science and mathematics education (Carey, 2000), in problem solving in computer programming (Davies, 1994), and in the difference between novice and expert knowledge in physics (Galili, Bendall, & Goldberg, 1993). Other research, within a developmental framework, has explored the ability of children of various ages to acquire new strategies in the laboratory (e.g., Alibali, 1999; Dixon & Bangert, 2002).

Although this research has indubitably established the presence of restructuring during naturalistic learning and human development, the accompanying theories are characterized by considerable diversity (e.g., Demetriou & Raftopoulos, 1999; Posner, Strike, Hewson, & Gertzog, 1982; Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001; Vosniadou & Brewer, 1987). For now, we focus on the framework by Posner et al. (1982) because it identifies explicit prerequisites for restructuring, before returning to naturalistic knowledge restructuring and theoretical alternatives in the General Discussion. According to Posner et al., knowledge restructuring occurs (a) when there is dissatisfaction with existing knowledge and (b) when a comprehensible alternative conception is known and available that also (c) is consistent with other existing knowledge and (d) appears productive. The first of these prerequisites, that people must be dissatisfied with their current knowledge before they restructure, turns out to be crucial in accommodating the somewhat conflicting results in category learning.[1]

Knowledge Restructuring in the Laboratory: Category Learning

We focus on categorization for several reasons. First, category learning is relevant to naturalistic settings: Not only is there a link between categorization and the acquisition of expertise (e.g., Medin & Edelson, 1988), but also, some naturalistic tasks are exact analogs of categorization in the laboratory (e.g., physicians' diagnoses of skin disorders; Brooks, Norman, & Allen, 1991). Second, and most important in the present context, categorization tasks can be designed to be solvable by several empirically identifiable alternative strategies.

There are several ways in which people's categorization strategies can be identified. Bourne, Healy, Parker, and Rickard (1999) asked participants after each category learning trial to indicate how they had performed the classification (e.g., by guessing, using a rule, or remembering the exemplar from an earlier trial). Bourne et al. found that most people transitioned at least once from one strategy to another (most often from rule- to exemplar-based responding), thus providing some indication of restructuring. Dixon and Bangert (2002) used a similar methodology in a categorization-like problem-solving task and also found that people shifted between overt strategies.

In the preceding studies, strategy use was determined on the basis of participants' protocols (Dixon & Bangert, 2002) or introspective identification and report (Bourne et al., 1999). Whereas these procedures have the advantage of fine temporal resolution (i.e., trial-by-trial identification of strategies), it is often preferable to have independent evidence for the nature of people's strategies that does not rely on introspection or self-report. In categorization, this can be achieved with properly designed transfer items (e.g., Medin, Altom, Edelson, & Freko, 1982), in particular, test items that constitute exceptions to one possible strategy but not to another. In one of Medin et al.'s (1982, Experiment 4) classic studies, participants could classify stimuli on the basis of any one of two single-dimensional rules with 75% accuracy, whereas the correlation between two other binary-valued features could support perfect performance. Specifically, a stimulus belonged to one category when two of its features had the same value, and it belonged to another when the values on those features were different from each other (i.e., an XOR relationship). Medin et al. found that after completion of training, people were sensitive to the XOR structure in the stimuli when classifying novel items. In an attempt to understand how the XOR strategy developed, McKinley and Nosofsky (1993, as cited in Nosofsky, Palmeri, & McKinley, 1994) replicated and extended the Medin et al. (1982) experiment. At various points in training, participants completed a transfer test that included novel items that conformed to the XOR rule but that were exceptions to the single-dimensional rule. Early in training, people applied the expedient single-dimensional rule, as revealed by low performance on those exception items. Later in training, people were found to have shifted to the more complex XOR strategy, as revealed by accurate categorization of the exception items.

In a related study, Johansen and Palmeri (2002) trained people to categorize a small number (9 or 10, depending on the experiment) of four-dimensional stimuli over the course of 32 training blocks. Responses to novel transfer stimuli, presented on six occasions throughout training, were used to determine each participant's representation and strategy. Learners initially used a rule-like representation that relied on only one stimulus dimension (out of four), which permitted imperfect classification. Later in training, most participants appeared to be using an exemplar-based representation that involved all dimensions.

The reports of representational change by Bourne et al. (1999), Dixon and Bangert (2002), Johansen and Palmeri (2002), and McKinley and Nosofsky (1993, as cited in Nosofsky et al., 1994) stand in contrast to situations in which people were found to resist knowledge restructuring.

[1] Although there is much consensus that "dissatisfaction" or performance error is a crucial determinant of knowledge restructuring, in some cases, restructuring has been observed in the absence of error (e.g., Dixon & Bangert, 2002; see also Dixon & Dohn, 2003). We defer discussion of those cases until the General Discussion.

Resistance to Restructuring in Category Learning

Lewandowsky et al. (2000) used stimuli that could be classified by several strategies. The stimuli varied along three dimensions, two of which were quasi-continuous and one that was dichotomous. The dichotomous dimension correctly predicted category membership 75% of the time. The two quasi-continuous dimensions were perfectly predictive when considered together, but the category boundary defined by their relationship was nonlinear and moderately complex. Unlike the preceding experiments, Lewandowsky et al. (2000) explicitly encouraged knowledge restructuring by revealing additional information about the task at some point in learning. When the information was revealed at the outset of training, participants adopted a complex strategy involving the two quasi-continuous dimensions, in preference to an expedient strategy involving the single dichotomous dimension that was learned by all other participants (identified by the misclassification of items that were exceptions to the expedient strategy). However, when the same extended information was revealed after participants had already adopted the expedient strategy, no strategy shift was observed.

The persistence of an expedient strategy in the presence of information that succeeded in encouraging use of a more sophisticated alternative when presented at the outset is counterintuitive and cannot be readily reconciled with the results discussed earlier. Why are some category learning tasks accompanied by spontaneous restructuring (e.g., Bourne et al., 1999; Johansen & Palmeri, 2002; Medin et al., 1982), whereas restructuring is resisted in others despite experimental encouragement (Lewandowsky et al., 2000)? We focus on performance error and the availability of information about an alternative strategy as possible points of reconciliation.

Error and Availability of an Alternative as Unifying Factors

The importance of error, or dissatisfaction of the learner, in longitudinal restructuring is well known (e.g., Posner et al., 1982). In category learning, the concept of error likewise permits an account of the results discussed so far, both positive and negative. Concerning a positive instance of restructuring, the role of error in the Johansen and Palmeri (2002) study is straightforward. Participants in that experiment could achieve perfect performance only by attending to all four stimulus dimensions and memorizing all instances (or by learning some complex rule, but few participants appeared to do so). Memorization was aided by the small number of training stimuli (at most 10), which rendered the availability of an alternative strategy (viz., memorization) obvious to participants, thus obviating the need for experimental encouragement.

In the negative case, restructuring was not observed, despite experimental encouragement, because error was low even with the initial expedient strategy. In the study by Lewandowsky et al. (2000), training commenced with two exception-free blocks during which the expedient dimension was perfectly predictive. Moreover, whereas the expedient dimension was binary and represented by verbal labels, the remaining numeric dimensions had many values. Because of the complexity of the alternative strategy, participants did not discover it unless they were explicitly instructed about it at the outset. In the absence of instructions, with the expedient dimension being highly salient and initially perfectly predictive, error was rapidly reduced, and the few exceptions (n = 4), when they occurred, were easily identified and memorized. Thus, even though the expedient dimension supported only 75% accuracy, participants performed above this level by memorizing the exceptions—this removed the need for restructuring and enabled people to ignore information about an alternative strategy when it was presented late in training.

We now present a series of experiments that examined this two-factor candidate account by systematically manipulating performance error and the availability of information about the complex alternative. Error was manipulated by altering the predictive value of an expedient single-dimensional rule or by changing the memorizability of exemplars that constituted exceptions to that rule. In all experiments, restructuring was experimentally encouraged by providing additional information about the task either before or during learning or not at all.

Category Learning Experiments

All experiments involved two phases, each including a set of training trials followed by a transfer test. Training stimuli were designed so that they could be classified by at least two strategies, one expedient (i.e., based on a single dimension of varying validity) and one complex (i.e., involving two dimensions that jointly permitted perfect performance). Several transfer items constituted exceptions to the expedient strategy (but not to the complex alternative) and thus would be classified in opposing ways depending on which strategy people used.

Experiment 1 used a single expedient dimension of moderately high validity. Exceptions to that dimension could not be memorized, and knowledge restructuring was observed when information about the complex alternative was revealed. When that information was withheld, no restructuring occurred. Experiment 2 manipulated the validity of the expedient dimension and established a direct link between the amount of performance error and extent of restructuring in response to information about the alternative strategy. Experiment 3 again used an expedient dimension of moderately high validity but permitted memorization of exceptional instances. No restructuring was observed when the alternative strategy was revealed.

Experiment 1

The first experiment investigated whether knowledge restructuring could be induced, in the presence of performance error, by revealing additional information about the task during learning. The expedient strategy (involving a single dichotomous stimulus dimension) permitted imperfect (i.e., 80% accurate) classification performance. In contrast to relevant previous studies (e.g., Lewandowsky et al., 2000; Medin et al., 1982), memorization of exceptions to the simple strategy was precluded here by presenting each exceptional stimulus once only. The complex alternative strategy depended on the nonlinear combination of two other quasi-continuous stimulus dimensions but permitted perfect classification.

Information about the complex strategy was either never presented (control condition) or was revealed in between training phases (reveal condition). Knowledge restructuring was examined by inferring people's strategies before and after the reveal manipulation on the basis of their transfer responses. The experiment also included partial transfer items that probed what people had learned about the diagnosticity of individual dimensions.

Method

Participants. The participants were 28 undergraduate volunteers from the University of Western Australia. Participants received credit toward completion of an introductory psychology class. Participants were randomly assigned to condition, with 13 participants in the reveal condition and 15 in the control condition. Both conditions involved two identical phases, each consisting of a sequence of training trials followed by a transfer test. The only difference between conditions was that, in the reveal condition, additional information about the task was made available in between phases.

Stimuli and apparatus. The stimuli were presented and controlled with a computer. The three dimensions appeared in the same screen position on every trial: The dichotomous expedient dimension (z, which we call the "label") appeared at the top, followed by the two quasi-continuous numeric dimensions (x and y, which together specify what we refer to as "location"). Responses were made with the arrow keys, with the left key denoting Category A and the right key denoting Category B. Both alternatives and their representative keys were displayed on the screen during each trial. The dimensions were instantiated as the type (z), size (x), and number of cars consumed (y) by a fictitious monster, and Categories A and B were denoted as "feels good" and "feels bad," respectively.

The category space with the distribution of training and transfer items is presented in Figure 1. During training, x and y each took on 12 values, resulting in 144 possible training items. All training items could be categorized correctly using two linear boundaries:

$$Category = \begin{cases} A \text{ if } 169 + 1.1x > y > 76 + 1.1x \\ B \text{ otherwise} \end{cases} . \qquad (1)$$

The boundaries ensured that the values of x were individually uninformative and that an equal number (77) of the 144 possible training stimuli fell within each category. Values of y were potentially informative on their own, and use of this single dimension could result in 70.8% correct performance (participants would have to use both y and y^2 optimally; y alone was nonpredictive).

The complex strategy relied on application of the bilinear boundaries defined by Equation 1 and could be verbally instantiated as follows: "If a monster has eaten enough but not too much for its size, then it feels good. If a monster has eaten very little or way too much for its size, it feels bad. Bigger monsters need to eat more to feel good."

The expedient but imperfect predictor was created by mapping each value of z onto a category, such that if $z = 0$ (one type of monster), the category was A with probability .80 and B otherwise (and, conversely, for $z = 1$, the other monster). Accordingly, on 115 out of the 144 training items, the label (1 or 0 with equal probability) predicted the category: $P(A|z = 0) = P(B|z = 1) = .80$. A verbal summary of the expedient strategy might be as follows: "Monsters of this type nearly always feel bad, whereas monsters of that type nearly always feel good."

People were said to rely on the expedient strategy if they classified stimuli on the basis of the label (z). Because each training item was presented only once in each phase, it was unlikely that participants could memorize the 29 "exceptional" exemplars (e.g., those for which the correct category was A, despite $z = 1$), all of which involved different locations (i.e., values of x and y).

Transfer stimuli (identified by triangles in Figure 1) were constructed separately, with 16 unique locations presented twice, once with each label, yielding a total of 32 transfer trials. All transfer stimuli were the same for all participants.

In addition, a set of "partial" transfer items was generated that contained only one of the stimulus dimensions. The partial items consisted of x only ($x = 25, 40, 55, 75, 95$), y only ($y = 124, 164, 212, 244, 276$) or z only ($z = 0, 1$). The values of x and y for partial stimuli were chosen such that they did not appear in any other training or transfer stimulus.

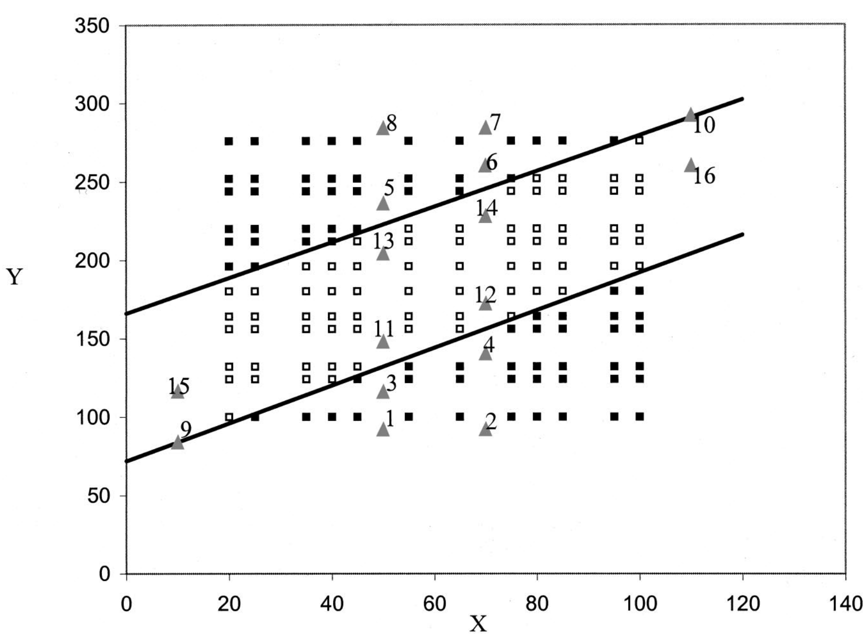

Figure 1. The stimulus space. Training instances are square markers, and category membership is shown by shading (open items are Category A, filled items are Category B): Numbers identify transfer stimuli (light-shaded triangles).

Procedure. Regardless of condition, the training sequence in each phase involved one presentation of each of the 144 training items. After a response was made, feedback (the word *correct* or the word *wrong*) was presented at the bottom of the screen for 2 s. The next trial began following a 1-s blank interval. The order of stimuli was randomized separately for each participant and phase.

The transfer sequence within each phase commenced with presentation of the 32 complete transfer items, one at a time in a new random order for each participant and phase. No feedback was given after transfer responses. The 12 partial transfer items were presented next. For each partial item, participants judged its likely category membership by positioning a cursor along a horizontal axis whose left and right endpoints reflected 100% likelihood that the stimulus would belong to A or B, respectively. Responses were coded from 1 (100% judged likelihood of A) to 32 (100% judged likelihood of B).

Participants in the reveal condition were informed of the complex strategy after the first transfer test and before the second learning sequence commenced. This was achieved with a graphical display that drew an analogy between the categorization problem and normal digestion, noting that large creatures need to eat more to feel good but that overeating makes one feel bad. Other than this, the display made no explicit mention of two boundaries and no depiction of the category space. The display was provided on a sheet of paper, was discussed briefly with the experimenter, and remained visible for the remainder of the session. In the control condition, no hints were provided, and participants were merely asked to learn the category structure as best they could.

Results and Discussion

Overview. Analyses of all studies revolved around the detection of knowledge restructuring. Accordingly, for the training data, we focused on performance on the diagnostic exception items. Poor performance on the exceptions would suggest that people used the expedient strategy, whereas good performance on the exceptions was indicative of some other strategy. By implication, if the reveal manipulation were followed by a pronounced improvement in performance on exceptions, this would be indicative of knowledge restructuring.

For the transfer data, we focused on people's sensitivity to the label, as revealed by how likely they were to switch classification responses for stimuli at identical locations presented with one or the other label. The more likely people were to alter categorization responses when the label was changed, the more they were relying on the expedient strategy. By implication, any reduction in the likelihood of switching between phases indicated the presence of knowledge restructuring.

Training. Each training sequence was divided into 4 blocks of 36 trials. Proportions correct for each type of stimulus (expedient rule: consistent vs. exceptional) are shown in Figure 2. The data were analyzed with a 2 (condition) × 2 (phase) × 4 (block) × 2 (type: rule consistent vs. exceptional stimulus) between–within analysis of variance (ANOVA). For this and all other analyses, we set a significance level of .05 and do not report exact p values.

As expected, rule-consistent items were generally more likely to be categorized correctly than were exception items, $F(1, 26) = 89.88$, $MSE = 0.260$, and participants improved overall during the experiment, as evidenced by the effect of phase, $F(1, 26) = 24.52$, $MSE = 0.016$. Those main effects were qualified by two interactions: First, participants in the reveal condition improved more in Phase 2 than did the participants in the control condition, resulting in a significant Condition × Phase interaction, $F(1, 26) = 14.08$, $MSE = 0.016$. Second, the disadvantage for the exceptional items was significantly reduced for the reveal condition in Phase 2, resulting in a three-way Type × Condition × Phase interaction, $F(1, 26) = 5.16$, $MSE = 0.096$. Taken together, the interactions are indicative of the presence of knowledge restructuring because the reveal condition benefited more from further training than did the control condition, and that benefit was particularly pronounced for the exception items.

Transfer. The analysis considered responses to individual transfer items to permit identification of each participant's strategy. Figure 3 shows the proportion of participants in both conditions and phases who classified each of the 32 stimuli as belonging

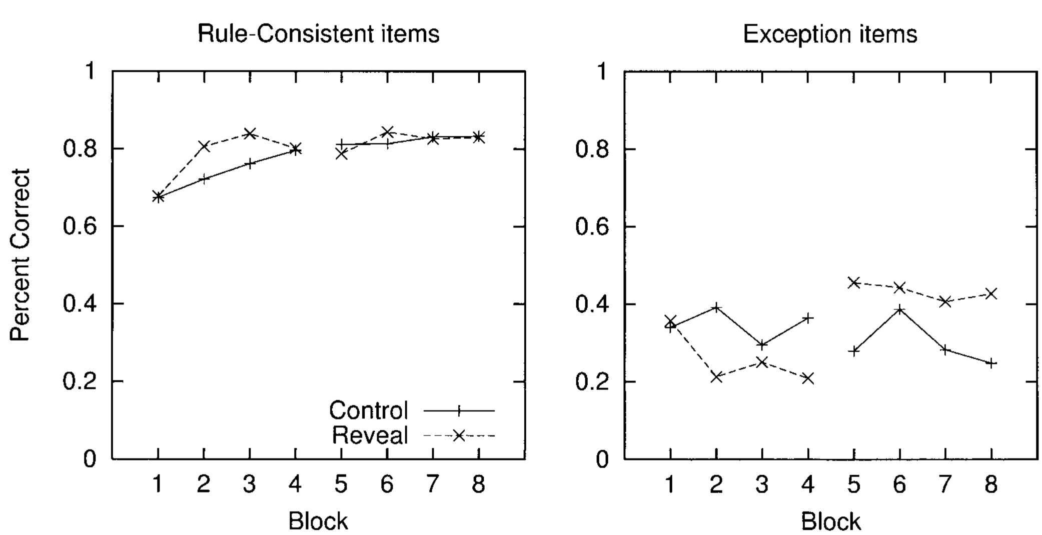

Figure 2. Experiment 1 proportion correct during training. Rule-consistent and exception items are shown separately. Blocks 1–4 are in Phase 1, and Blocks 5–8 are in Phase 2.

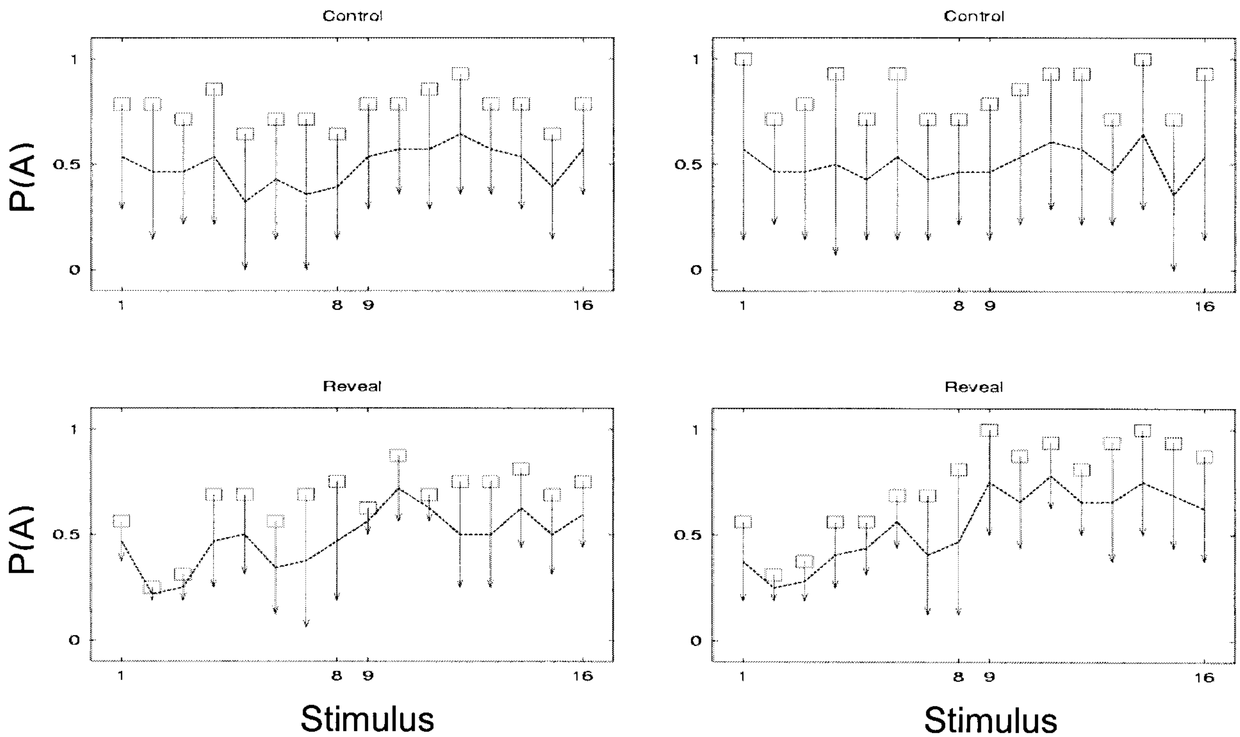

Figure 3. Experiment 1 probability of Response A for each transfer stimulus. The abscissa corresponds to the stimulus location numbers shown in Figure 1. Each stimulus location was presented twice, once with each of the two labels. The tail of each vector is the mean response to one label (associated with Response A), and the head the response to the other. Vector lengths represent the effect of the label. Vectors are connected at their means by a dashed line representing the effect of stimulus location. The top row is the control condition, the bottom is the reveal condition, the left column is Phase 1, and the right column is Phase 2.

to Category A. The abscissa identifies the transfer stimuli (i.e., each *x, y* location) according to the numbering in Figure 1. According to the bilinear boundaries used for stimulus construction, the eight locations on the left belong to Category B, and the remaining eight items belong to Category A. If people successfully apply the complex strategy, their response profiles are expected to follow a step function in the figure, with probabilities near zero on the left and near unity on the right.

The figure also reveals people's sensitivity to the label by showing two responses for each location connected by arrows. The tail of each arrow represents the response proportion when the label was zero ($z = 0$) and the head the proportion when $z = 1$. The lengths of the arrows in each panel thus provide a direct visual impression of the extent to which people relied on the expedient strategy.

Visually, knowledge restructuring would be identified by long arrows in Phase 1 that are vertically aligned, and short arrows in Phase 2 that are vertically offset in two clusters (close to zero on the left and close to unity on the right). The dashed line connecting the center points of the arrows represents the average response across labels for each location and permits ready visual detection of people's sensitivity to location.

Figure 3 shows that participants in both conditions responded quite similarly at the end of Phase 1 but quite differently following Phase 2. A between-within ANOVA involving the variables condition, phase, label, and location showed that the differences in Phase 2 were reflected in a Condition × Phase × Label interaction, $F(1, 26) = 11.97$, $MSE = 0.47$. This interaction reflects the fact that, despite nearly identical use of the label in Phase 1 (where, if anything, the reveal group used the label more than the control group, as shown by their longer arrows), in Phase 2, the reveal group's arrows were shorter. A significant interaction contrast for Condition × Label for Phase 2, $F(1, 26) = 6.92$, $MSE = 0.036$, confirms that the reveal group used the label significantly less than did the control group.

The reduced use of the label by the reveal condition in Phase 2 was accompanied by an increased use of stimulus location, shown in Figure 3 by the deviation of the dashed line from the horizontal. This effect of the reveal manipulation was reflected in a significant omnibus Condition × Phase × Location interaction, $F(15, 390) = 2.36$, $MSE = 0.10$, and is confirmed by a significant Phase × Location interaction contrast for the reveal group alone, $F(15, 180) = 3.31$, $MSE = 0.66$. Visually, the dashed line connecting the centers of the arrows for that condition was transformed from being near 0.5 across all locations in Phase 1 to something closer to a step function in Phase 2, with low probabilities on the left and higher probabilities on the right.

The clear change in strategy observed in the reveal condition stands in striking contrast to the absence of any such change in the control condition, indicated by the fact that the dashed line remained relatively flat in both phases (i.e., a nonsignificant Phase × Location contrast for this group, $F[15, 210] = 1.05$, $MSE = 0.75$) and that the lengths of the arrows did not change between phases (i.e., a nonsignificant Phase × Label interaction contrast for this group, $F[1, 14] = 3.37$, $MSE = 0.68$). Taken together, the results of both conditions confirm that knowledge restructuring occurred in Experiment 1 but that it arose only as a consequence of the reveal manipulation rather than as an inevitable by-product of further training.

Partial transfer items: Contingency ratings. The partial stimuli provide information about the learned strength of each dimension on its own. Table 1 shows that, on average, participants in both conditions knew that only the label was individually predictive of category membership. This was confirmed by three separate three-way ANOVAs—one for each dimension—involving condition, phase, and the value of the partial stimulus. For the label, there was a significant change in response as the value went from 0 to 1, $F(1, 26) = 149.6$, $MSE = 79.30$. No other significant differences emerged, and the critical three-way Condition × Phase × Value interaction was not significant, $F(1, 26) = 1.63$, $MSE = 31.01$. This shows that people's judged contingency for the label remained unchanged across phases in all conditions, notwithstanding the fact that knowledge restructuring was unambiguously identified by the classification responses in the reveal condition. Knowledge restructuring therefore did not entail the unlearning of the strong contingency between label and category but instead consisted of people choosing to ignore a contingency that they had previously relied on. We return to this dissociation between contingency ratings and classification responses after the remaining data have been presented.

For the y dimension, value entered into a main effect, $F(4, 104) = 3.15$, $MSE = 47.62$, and no other effects were significant; the critical three-way interaction had $F(4, 104) < 1$. Participants were particularly insensitive to the x dimension; there were no effects involving its value, with the main effect and three-way interaction Fs < 1. Thus, participants were sensitive to both of the partially valid dimensions, y and the label, and that sensitivity did not change appreciably with condition or phase.

Experiment 2

Experiment 1 showed that revelation is a necessary condition for restructuring, whereas performance error, by itself, is insufficient. Experiment 2 further investigated the interplay between performance error and the revelation of information about a complex alternative. Experiment 2 controlled the extent of error by manipulating the validity of the expedient strategy (i.e., what proportion correct its optimal use could produce) from 60% through 100%.

Table 1
Contingency Judgments for Partial Stimuli in Experiment 1

Dimension	Stimulus 1	2	3	4	5
x	16	16	16	14	14
y	15	13	13	16	17
Label	25	5			

Note. Judgment scales ranged from 1 (Category A) to 32 (Category B). Table entries are averaged over conditions and phases. The values for each stimulus are given in the text.

Method

Participants. Sixty students from an introductory psychology course at the University of Western Australia volunteered for course credit. Twelve participants were randomly assigned to each of five conditions labeled R60, R80, R90, R100, and throughout (T60). In the first four conditions, the reveal manipulation was administered between training phases, similar to Experiment 1, with the number in the name referring to the validity of the label. The final condition (T60), presented information about the complex strategy at the outset of training (label validity 60%). The control condition from Experiment 1 was not included in this study.

Stimuli. The items were the same as those used in Experiment 1 except that the partial transfer items were omitted. For the R60 and T60 conditions, the validity of z was .60, such that for 87 of the 144 training items, the label (z) predicted the category: $P(A|z = 0) = P(B|z = 1) = .60$. Those 87 items were divided evenly between the two categories, with the extra invalid trial randomly assigned (e.g., 43 out of 72 Category A items had $z = 0$, whereas 44 of the 72 Category B items had $z = 1$). For the R80, R90, and R100 conditions, the label validities were .80 (114 or 116 out of 144 items as in Experiment 1), .90 (128 or 130 out of 144), or 1.0, respectively.

Procedure. The procedure was identical to that of Experiment 1, with the following exceptions: First, the 12 partial transfer items were not presented. Second, in the throughout condition (T60), the display was presented, and information about the complex strategy was revealed, before training commenced.

Results

Training. Aggregating trials into blocks in the same way as in Experiment 1, Figure 4 shows training performance separately for each type of stimuli. The figure shows that, whereas performance differed little between conditions for the rule-consistent items, there was considerable variation in the performance on exceptional items. Because the R100 group did not receive any exception items, it was excluded from the Condition (4) × Phase (2) × Block (4) × Type (2) ANOVA. The analysis revealed no significant effects involving block, although the main effects of condition, $F(3, 44) = 5.71$, $MSE = 0.10$, phase, $F(1, 44) = 47.40$, $MSE = 0.026$, and type, $F(1, 44) = 57.32$, $MSE = 0.216$, were all significant. The condition effect was due primarily to the strong performance of the higher validity groups, and the phase effect arose from the increased performance in Phase 2. The type effect reflected the fact that performance was better for rule-consistent items than for exception items. Critically, as in Experiment 1, these main effects were qualified by a Condition × Phase interaction, $F(3, 44) = 5.51$, $MSE = 0.026$, and a Condition × Phase × Type interaction, $F(3, 44) = 2.85$, $MSE = 0.069$. Because these interactions are potentially diagnostic of knowledge restructuring, they were explored further with a series of planned comparisons.

The difference between the R60 and T60 groups was examined with two Condition × Type ANOVAS, one for each phase. The R60 group was performing worse than the T60 group during Phase 1 on both types of items, which was confirmed by the main effect of condition, $F(1, 22) = 9.88$, $MSE = 0.007$. This result confirmed that the additional information about the complex strategy, when presented at the outset, assisted learning in the throughout condition. In addition, the two conditions were not identifiably different

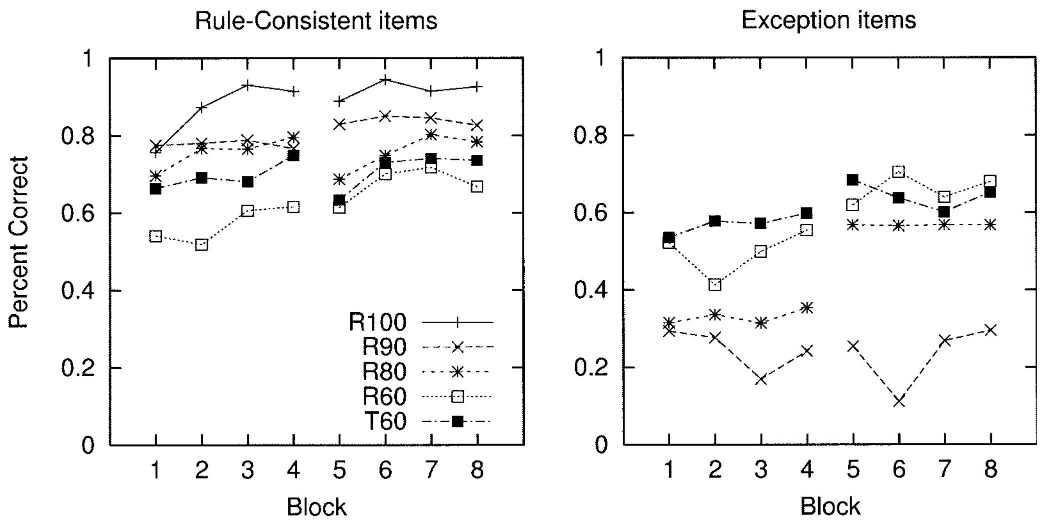

Figure 4. Experiment 2 proportion correct during training. Rule-consistent and exception items are shown separately. Blocks 1–4 are in Phase 1, and Blocks 5–8 are in Phase 2.

during Phase 2, $F(1, 22) < 1$; this in turn shows that the R60 condition benefited from the reveal manipulation after the first phase and was able to match the performance of the throughout condition from then on. Moreover, the combined fact that the R60 condition differed from T60 in Phase 1 but not in Phase 2 indicates that there was no spontaneous restructuring in the R60 group, similar to the persistent use of the expedient strategy in the control condition of Experiment 1. Hence, the presence of massive performance error, by itself, was insufficient to elicit restructuring. In addition to error, restructuring required experimental encouragement in the form of the reveal manipulation (the transfer analysis below affirms this conclusion).

Figure 4 also shows that the reveal manipulation had a large effect on the R80 group, dramatically increasing performance on the exception items while slightly decreasing performance on rule-consistent items. This is shown to be a significant difference by a Phase × Type interaction in a separate ANOVA for the R80 condition, $F(1, 11) = 8.29$, $MSE = 0.021$. The error rates on exception items for the R80 group change from a level near probability matching (approximately 33%) to one further away (55% correct). The error rates change less for the rule-consistent items, in part because performance there is above chance during Phase 1.

The R90 group, by contrast, did not exhibit any change in performance between phases, suggesting that the reveal manipulation remained ineffective for this level of validity. In Figure 4, this is apparent from the virtually unchanged performance on exceptions across phases. Statistically, the mean proportion correct on exceptions in Phases 1 and 2 were equivalent (0.267), so $F(1, 11) = 0$ for phase. The levels of performance in both phases, for both types of items, are close to the error rate one would expect if participants were probability matching throughout the experiment.

Transfer. Responses were analyzed in the same manner as for the first experiment and are shown in Figure 5. A 5 × 2 × 16 × 2 (Condition × Phase × Location × Label) between–within ANOVA revealed a significant three-way interaction involving label, condition, and phase, $F(4, 55) = 3.23$, $MSE = 0.25$. This overarching interaction reflects the selectively diminished effect of label (i.e., reduced use of the expedient strategy) as a result of the reveal manipulation following Phase 1, for the moderate validity conditions. Specifically, whereas participants in the R80 condition used the expedient strategy reliably in Phase 1 (Panel C; the arrows all point in the same direction), their responses in Phase 2 (Panel H; the arrows are significantly shorter and more variable in direction) were largely insensitive to the label. A similar effect, though smaller in magnitude (because the Phase 1 arrows are already small), was obtained for the R60 condition (Panels D and I).

The overall Label × Condition interaction, $F(4, 55) = 21.95$, $MSE = 1.09$, reflects the high reliance on the label in the R100 group and its reduction as validity decreased. There was also an overall Label × Phase interaction, $F(1, 55) = 17.35$, $MSE = 0.25$, reflecting the success of the reveal manipulation overall, as label use was reduced in Phase 2. However, these two interactions and the main effect of label, $F(1, 55) = 126.01$, $MSE = 1.09$, are of little interest in light of the above overarching three-way interaction.

The effect of location and, by implication, use of the complex strategy, can be ascertained from the profile formed by the dashed lines in Figure 5. The three-way interaction of location, condition, and phase, $F(4, 55) = 3.23$, $MSE = 0.25$, confirms that location plays a different role in different conditions and across phases. The Location × Condition interaction, $F(60, 825) = 2.69$, $MSE = 0.2$, and the Location × Phase interaction were also significant, $F(15, 825) = 3.46$, $MSE = 0.16$, as was the main effect of location, $F(15, 825) = 22.27$, $MSE = 0.22$. Overall, participants in the low-validity conditions used the stimulus location most, and participants in the moderate validities, especially in the R80 group, changed their response pattern from one dominated by the label (flat dashed lines and long arrows in Figure 5) to one dominated by location (approximately step-shaped dashed lines, with short arrows).

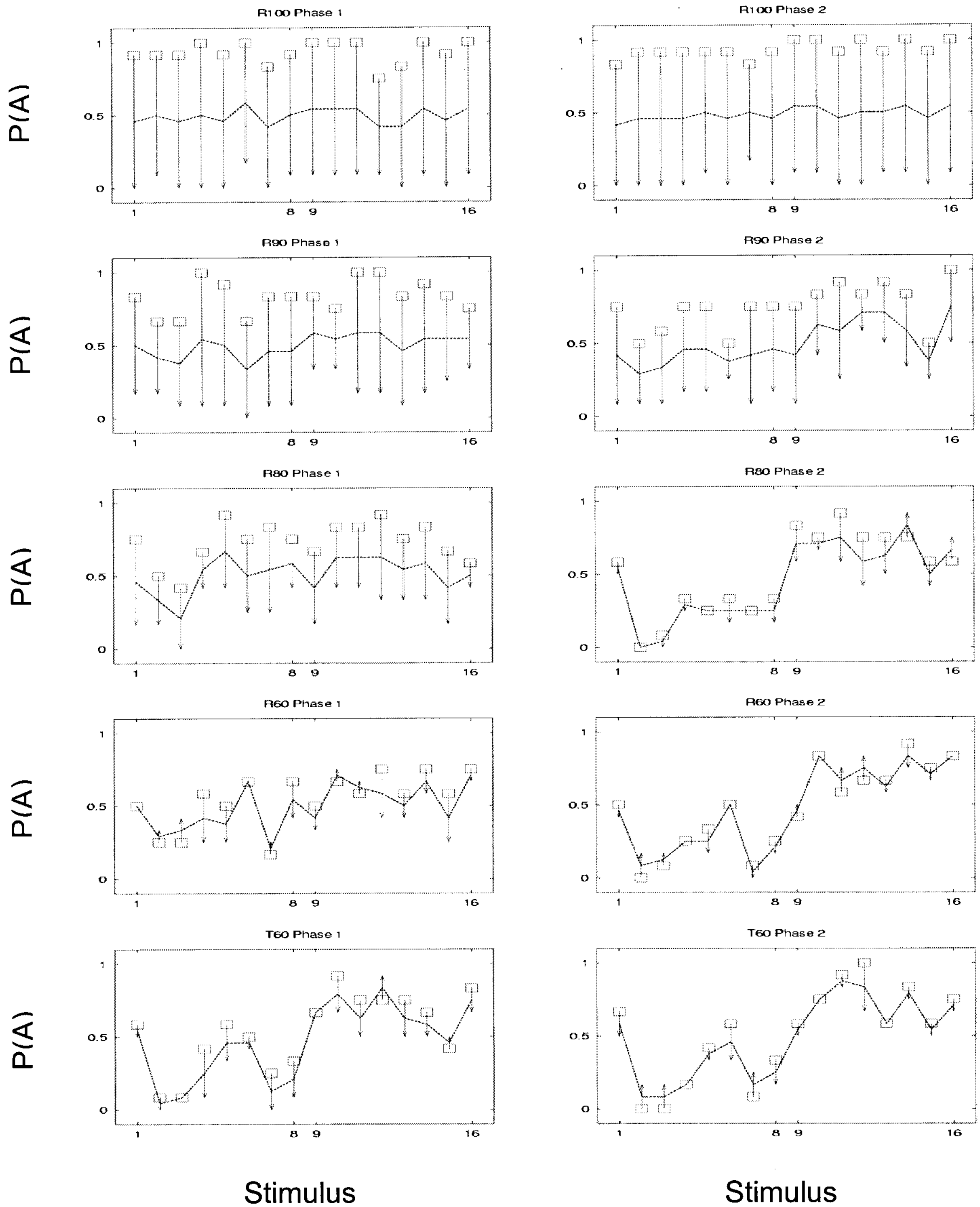

Figure 5. Experiment 2 probability of Response A for each transfer stimulus. The abscissa corresponds to the stimulus numbers shown in Figure 1. From top to bottom, the panels show results of the R100, R90, R80, R60 and throughout groups.

The minimal change in the responses of the R60 group raises the possibility that people in that condition either spontaneously restructured their knowledge during Phase 1 or never adopted the marginal (60% valid) expedient strategy in the first place. The earlier analysis of the training data speaks against this possibility because it identified clear differences between those two conditions in Phase 1 that disappeared in Phase 2. However, the transfer data are more ambiguous: A Condition × Label × Location interaction contrast on just these two groups at Phase 1 did not show any effects involving condition. We therefore performed a post hoc discriminant function analysis on these two groups for Phase 1.

Each participant gives two responses to each transfer stimulus location, one for each value of the label. We took as data the number of times out of 16 that a participant could switch his or her response, just as we did in Experiment 1 when investigating the role of error. These 16 scores were 100% successful in discriminating participants in the R60 group from those in the T60 group during Phase 1 but were unable to differentiate completely between the groups during Phase 2. Thus, it is clear that, despite similarities between the groups in terms of average response rates, the two groups are significantly different from one another during the first, but not the second, transfer phase.

Error and strategy change. We next examined what factors, experienced by an individual participant during training, might be related to the extent of that person's knowledge restructuring. As a measure of restructuring at the individual level, we pooled all responses in a given transfer phase and computed the number of locations, out of 16 possible, for which the participant's response switched from A to B or vice versa when the label changed between presentations of an item in the same location. We computed the correlation between error in the last block of training in Phase 1 and the change in label sensitivity (as captured by this switch score) between phases for each participant. We excluded the throughout condition, for which no restructuring was expected. For the 48 participants in the reveal conditions, the correlation between error rate and label-sensitivity difference was significant, $F(1, 46) = 4.66$, $SE_{\text{est}} = 3.86$, accounting for 9.2% of the variance in label sensitivity. Likewise, the condition variable (dummy coded for the regression) accounted for 15.5% of the variance, which is also a significant proportion, $F(3, 44) = 2.89$, $SE_{\text{est}} = 3.83$. These two predictors were significantly collinear; 50.6% of the variance in individual error scores is accounted for by the validity of the label.

Discussion

The results of Experiment 2 are readily summarized: When a single dimension was at least moderately predictive of category membership during initial training, participants learned to rely on this expedient dimension. At the same time, as long as the expedient dimension was not highly valid, participants were responsive to a manipulation that pointed out an alternative, more complex two-dimensional strategy. In essence, when the single expedient dimension was more unsatisfactory than the revealed alternative, participants complied with the recommendation to restructure their approach to the task. This result establishes performance error as a necessary condition for knowledge restructuring in categorization. The earlier result from Experiment 1, that restructuring does not occur spontaneously when additional information is withheld, establishes revelation as another necessary condition for restructuring, at least in situations in which the complexity of the alternative prevents participants from spontaneously discovering it. Experiments 1 and 2 thus identify error and revelation as individually necessary and jointly sufficient for knowledge restructuring.

The results of the first two experiments, and the conclusion we have just drawn, need to be reconciled with the divergent findings of Lewandowsky et al. (2000). In that study, restructuring was absent after a similar reveal manipulation, even though the validity of the expedient dimension was 0.75—exactly within the range in which a shift to a more complex strategy was observed in the present experiments.

There are several differences between the two sets of studies that might have given rise to the different outcomes. Within the error framework, the most critical of these differences should be the fact that Lewandowsky et al. (2000) used a small number of frequently repeated stimuli, whereas the present studies contained no repetition within a phase. Thus, although the validity of the expedient dimension was comparable, validity was not the sole determinant of the expedient strategy in the study by Lewandowsky et al. Instead, their expedient strategy involved the expedient dimension *and* memory for the few (4) exceptions that were mispredicted by it.[2]

In our Experiments 1 and 2, memorization was unavailable to assist with imperfect expedient rules. It follows that people should no longer restructure their knowledge when encouraged to do so if exceptions can be memorized. The final experiment explored this possibility.

Experiment 3

In this experiment, only a single exception stimulus was used in each category that was presented repeatedly within each phase. We again included the contingency rating task used in Experiment 1 with an additional set of two-dimensional partial stimuli.

Method

Participants. Forty-six students from an introductory psychology class volunteered for course credit. Participants were randomly assigned to one of three conditions: reveal ($n = 16$), control ($n = 14$), or throughout ($n = 16$). Participants in those conditions received information about the complex strategy in the middle of training, not at all, and at the outset, respectively.

Stimuli and procedure. Stimuli were constructed as in the first two experiments, with two exceptions. First, 116 items (instead of 144) were randomly sampled for each participant. The label was randomly assigned such that 114 items were correctly classified by that dimension (57 from each category), whereas the remaining 2 items (1 from each category) constituted exceptions. Because each of the exceptions was repeated 15 times, thus yielding the usual total number of training trials (144), the validity of the expedient dimension was 80% in all conditions.

Second, in addition to the single-dimensional partial stimuli used in Experiment 1, a further set of eight two-dimensional partial stimuli were created that presented a conjunction of x and y values for a contingency judgment. The chosen stimuli were (10, 84), (10, 116), (70, 140), (50, 148), (70, 228), (50, 236), (110, 260), and (110, 292), which are the locations of 8 of the 16 transfer items. In all other respects, the procedure was identical to that of Experiment 1.

Results and Discussion

Training. The training data showed a steady increase in performance across all conditions in both phases (see Figure 6). A 3

[2] The present experiments also differ from those of Lewandowsky et al. (2000) by not including two training blocks at the outset of each phase during which the exceptional items were withheld.

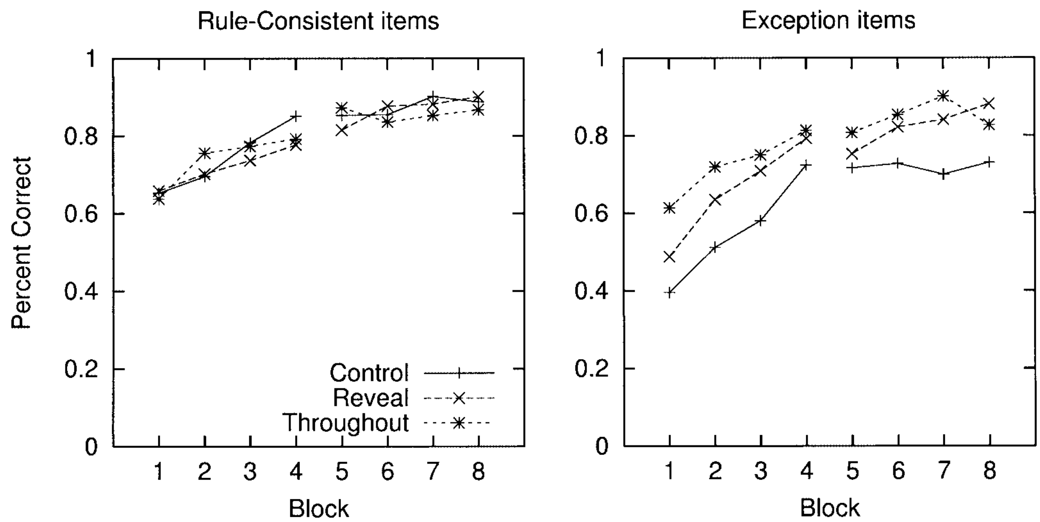

Figure 6. Experiment 3 proportion correct during training. Rule-consistent and exception items are shown separately. Blocks 1–4 are in Phase 1, and Blocks 5–8 are in Phase 2.

(Condition) × 2 (Phase) × 4 (Block) × 2 (Type) ANOVA showed significant effects for phase, type, and block. The main effect of phase was due to better performance in Phase 2 than in Phase 1, $F(1, 43) = 83.18$, $MSE = 0.045$. Both the main effect of block, $F(3, 129) = 31.41$, $MSE = 0.018$, and the Block × Phase interaction, $F(3, 129) = 11.02$, $MSE = 0.020$, stemmed from the rapid performance improvement in Phase 1 relative to the slower learning in Phase 2.

There was also a marginal effect of type, with performance being slightly better overall for rule-consistent items ($M = .80$) than for exceptions ($M = .73$), $F(1, 43) = 4.57$, $MSE = 0.234$. It is noteworthy that there was no sign of any effect involving condition, including the crucial three-way Condition × Phase × Type interaction, $F(2, 43) < 1$, indicating that no significant change in learning took place when the complex alternative was revealed. Comparison of Figure 6 with the training data of the first two studies (Figures 2 and 4) confirms that people's responses to the two unique exceptions in Experiment 3 were far more accurate than were the corresponding responses in the earlier studies. There is every reason to suggest that this reflects the opportunity for memorization afforded by the frequent repetition of the two unique exceptional stimuli. It is not possible to ascertain from the training data alone whether the lack of any difference between the throughout condition and the other two groups on the exception items in Phase 1 is due to all participants sharing a single strategy (memorization of exceptions) or to different groups using different response strategies.

Transfer. The transfer data, shown in Figure 7, provided no indication of restructuring but identified a difference between the throughout condition and the remaining two groups. Confirming the lack of restructuring, the 3 (Condition) × 2 (Phase) × 16 (Location) × 2 (Label) ANOVA revealed no significant interactions involving phase and condition. Instead, the overall differences between groups that are apparent in the figure were reflected in two interactions. First, the Condition × Location interaction, $F(30, 645) = 1.50$, $MSE = 0.22$, represented the greater reliance on the complex strategy by participants in the throughout group. We confirmed this in a post hoc test by combining the control and reveal groups and comparing their use of the location information to that by the throughout group; this Group × Location contrast was significant, $F(15, 660) = 2.62$, $MSE = 0.21$. Second, the Condition × Location × Label interaction, $F(30, 645) = 1.70$, $MSE = 0.11$, reflected the fact that the throughout condition not only used the location information more than did either of the other two groups but also used the expedient strategy less. The post hoc Group × Label contrast confirmed this, $F(1, 44) = 4.68$, $MSE = 2.12$. The main effects of both location and label were robust, $F(15, 645) = 6.095$, $MSE = 0.22$, and $F(1, 43) = 96.52$, $MSE = 2.16$, respectively, but are of little interest given the qualifying overarching interactions. Finally, label interacted with phase, $F(1, 43) = 4.57$, $MSE = 0.38$, reflecting the greater use of the label in Phase 2 compared with Phase 1; a post hoc test showed that this difference was significant for the pooled control and reveal groups, $F(1, 29) = 5.30$, $MSE = 0.32$, but not for the throughout group, $F(1, 15) < 1.0$.

Overall, the transfer data mesh with the results of Lewandowsky et al. (2000): Participants were able to use information about the complex strategy when it was presented at the outset, but they were unable, or unwilling, to use the information to restructure their knowledge after they had learned to rely on an expedient strategy. By implication, the knowledge restructuring observed in Experiments 1 and 2 arose from the error associated with the inability to memorize exceptions (or indeed any other training stimuli).

Contingency ratings. Analyses of ratings of the single-dimension stimuli (shown in Table 2) were carried out with three separate 3 (Condition) × 2 (Phase) × Value (ns = 2, 5, 5 for the values of z, x, and y, respectively) ANOVAs. People were found to be sensitive to the value of z, $F(1, 43) = 53.10$, $MSE = 153.60$, and y, $F(4, 172) = 2.97$, $MSE = 41.06$, but not x, $F(4, 172) = 1.80$, $MSE = 39.32$. None of the dimensions revealed an effect that depended on phase, and condition entered into only one marginal

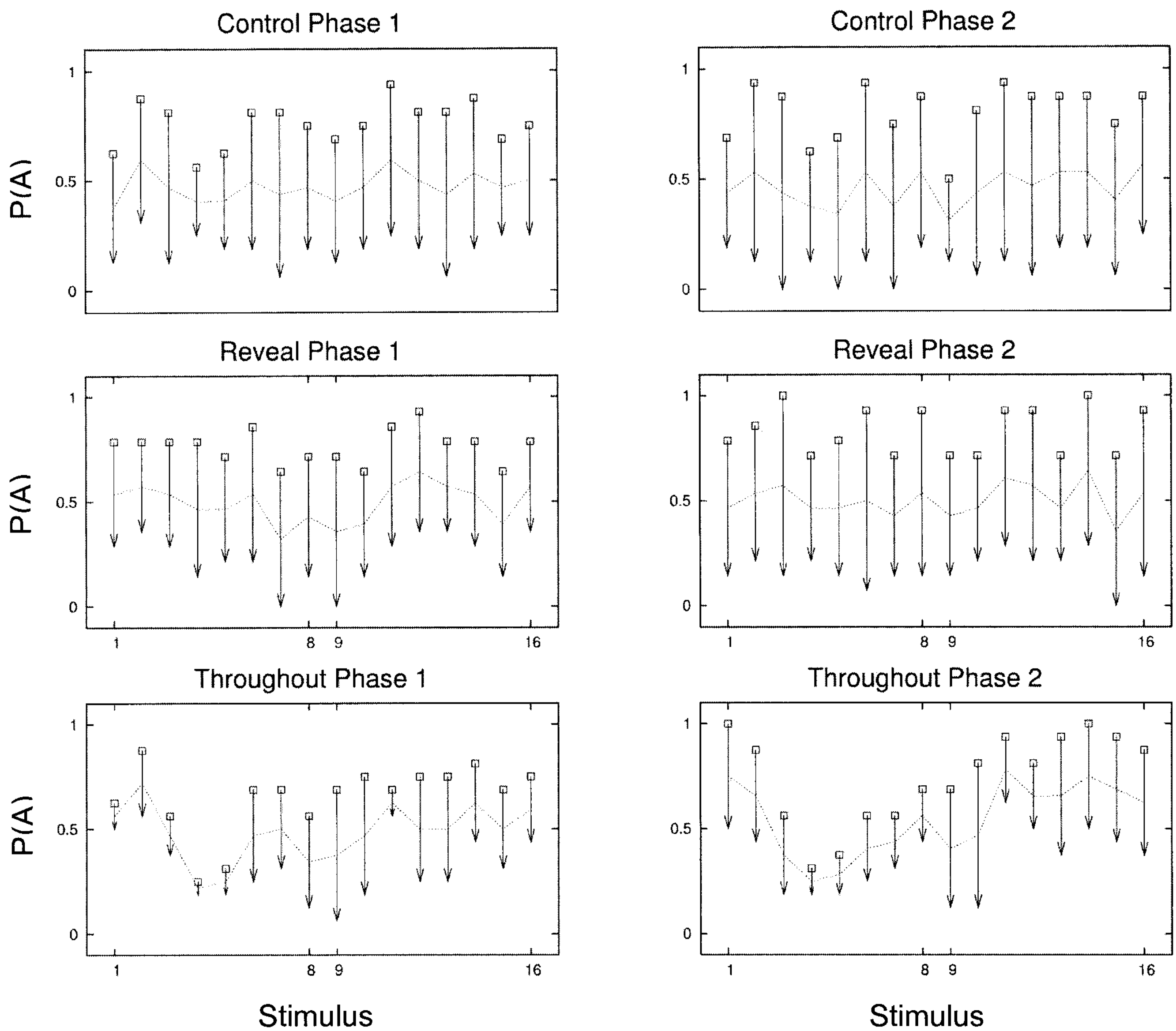

Figure 7. Experiment 3 probability of Category A response for each transfer stimulus. The abscissa corresponds to the numbers shown in Figure 1; rows are control, throughout, and reveal, from top to bottom.

interaction, namely, Condition × Value for *x*, $F(8, 172) = 2.00$, $MSE = 39.32$.

Thus, the contingency ratings confirm that participants developed an early understanding of the role of the label as a single predictor, understood that *x* was not individually predictive, and knew that *y* was only moderately predictive, and the revelation of additional information had little or no effect in any of the conditions. As in Experiment 1, the group differences observed in the classification of transfer items were not accompanied by differences in knowledge about individual dimensions.

Turning to the two-dimensional partial items, there was no significant interaction involving condition and phase and no significant effect involving phase (all Fs < 1). The location of the eight stimuli did seem to matter, however, with both a significant main effect of location, $F(7, 301) = 7.83$, $MSE = 93.47$, and a significant interaction of location and condition, $F(14, 301) = 1.82$, $MSE = 93.47$. The interaction reflects the slightly higher sensitivity to location by the throughout group (variance among eight responses was 14.0 for throughout, compared with 12.7 and 11.4 for control and reveal, respectively). This result supports and extends the analysis of the transfer and training data.

General Discussion

The results from our experiments converge on three principal conclusions that we discuss before dealing with a potential methodological limitation.

Table 2
Contingency Judgments for Single-Dimensional Stimuli in Experiment 3

	Stimulus				
Dimension	1	2	3	4	5
x	17	16	15	14	15
y	16	14	14	16	16
Label	22	9			

Note. Judgment scales ranged from 1 (Category A) to 32 (Category B). Table entries are averaged over conditions and phases. The values for each stimulus are given in the text.

Conclusion 1: The Need for an Available Alternative

As demonstrated by the control conditions in Experiments 1 and 3, the complex strategy was difficult to learn without further information. Even when the expedient strategy was of low validity (R60 condition of Experiment 2), people, after prolonged training, approached the task differently than did participants in the throughout condition. It follows that revelation is necessary for knowledge restructuring to occur in categorization. This conclusion is accompanied by two known exceptions.

First, other studies have observed spontaneous restructuring without experimental encouragement of any kind (Johansen & Palmeri, 2002; Medin et al., 1982). These results can be reconciled with ours if one recalls that the few stimuli used in these other studies were repeated many times, rendering memorization nearly inevitable. Because instance-based responding rapidly afforded perfect performance, people abandoned whatever expedient rules they may have used initially. In our experiments, by contrast, at most, two of the 144 training instances were repeated within a phase, thus precluding any shift to a strategy based on gradually emerging memorization. In support, Rouder and Ratcliff (2004) showed that people use instance-based categorization only if training stimuli are identifiable and memorizable (e.g., through repetition). We therefore suggest that spontaneous restructuring can occur in categorization if memorization of training instances is facilitated. When memorization is precluded, restructuring from one nonmemorial strategy to another requires experimental encouragement.

Second, in our studies people performed the task as expediently as possible, by relying on a single dimension. A corollary of this preference for expediency is that restructuring in our experiments always involved a shift to a more complex strategy. Shifts in the reverse direction, from complex to simple, may occur spontaneously. For example, in Logan's (1988) theory of skill acquisition, performance is thought to be driven by a complex algorithm at the outset, before people shift to a simpler, memory-based performance. During the acquisition of mental arithmetic, children likewise commence by counting on their fingers and then switch to a simpler, memory-based approach (Koshmider & Aschcraft, 1991; Shrager & Siegler, 1998).

Taken together, these two exceptions imply that knowledge restructuring can occur without experimental encouragement when people recognize the availability of an alternative strategy, either because it is simple or because there are relatively few instances to be memorized. In all other cases, we suggest that revelation is a necessary condition for restructuring.

Conclusion 2: The Need for Performance Error

Revelation by itself was insufficient to induce restructuring in any of our experiments unless performance error was also present. We observed restructuring in response to revelation when the validity of the expedient rule was low to moderate and memorization could not contribute to responding (Experiments 1 and 2). By contrast, we observed resistance to restructuring when the expedient rule was perfect (Experiment 2) or when exceptions to an imperfect expedient rule could be memorized (Experiment 3). It follows that the presence of performance error is a necessary condition for knowledge restructuring, whereas error and revelation are jointly sufficient.

Conclusion 3: The Persistence of Earlier Information

Third, as shown by the contingency ratings to partial stimuli (Experiment 1), restructuring is not accompanied by an unlearning of previously relevant contingencies between the expedient dimension and the response categories. People continue to be aware of the predictive role of the label, even though they discontinue its use during classification upon restructuring.

Limitation

The cover story for our stimuli and the reveal manipulation appealed to people's intuitions about digestion. This may have enabled people to use the information to shift away from the (intuitively meaningless) label to a strategy that was in accord with prior knowledge. At first glance, this may limit the applicability of our conclusions to situations in which the complex strategy is compatible with prior knowledge. However, this concern can be at least partially allayed by noting that the cover story used by Lewandowsky et al. (2000), which concerned the spread of wildfires, did not relate to people's prior knowledge. Lewandowsky et al. nonetheless found that people were able to learn a complex strategy if it was revealed at the outset but not after an expedient alternative had been acquired, which is exactly as we found here in Experiment 3. We therefore suggest that our results likely generalize beyond the particular stimuli used here. We now discuss the broader implications of our three conclusions.

Persistence of Original Knowledge After Restructuring

In naturalistic settings, the fate of prior knowledge after restructuring has been vigorously debated. Do people simply accumulate new knowledge (Wellman & Gelman, 1992), or do they build entirely new cognitive structures (Carey, 1985)? If they build new structures, do they maintain (Johnson & Carey, 1998) or discard (Novak, 2002) old structures that are no longer appropriate?

The evidence reported here suggests that people continue to have full access to their original knowledge. In Experiments 1 and 3, people were able to use the expedient dimension when presented on its own (in contingency judgments), irrespective of which strategy was used to classify the transfer stimuli proper. Thus, even when people demonstrably ignored the label during classification and instead relied on the complex alternative (e.g., Phase 2 of the reveal condition in Experiment 1), their categorization judgments to single-dimensional stimuli exhibited the same sensitivity to the label as did those of participants whose knowledge had not been restructured (e.g., Phase 2 of control in Experiment 2 or Phase 2 of reveal in Experiment 3). This suggests that people did not unlearn their expedient knowledge but instead chose not to use it after restructuring.

The persistence of old information is compatible with several theories of skill acquisition, which similarly assume that old strategies remain available even when they are no longer being used. For example, Logan's (1988) instance theory postulates that performance involves a race between a slow but accurate algorithm and memorized instances of previous problem solutions. Even though performance is eventually almost exclusively driven by instances, the original algorithm is assumed to remain intact and fully operational.

Turning to another domain, there is much evidence that children repeatedly restructure their knowledge while learning to perform single-digit mental arithmetic (e.g., Shrager & Siegler, 1998; Siegler, 1987). This research has revealed two attributes of developmental restructuring that are relevant here. First, there is a clear chronological hierarchy of strategies, with some strategies (e.g., counting both addends using fingers) emerging before others (e.g., counting from first addend), which in turn are followed by even more efficient shortcuts (e.g., retrieving the answer from memory). Second, notwithstanding this nearly invariant order of discovery, there is no evidence that an early strategy is unlearned after a newer one has been discovered. On the contrary, some techniques—such as counting fingers versus retrieving the answer from memory—may coexist for several years after discovery of the latest strategy (e.g., Shrager & Siegler, 1998).

Finally, there are intriguing instances in which the disappearance of earlier knowledge after further learning turns out to be more apparent than real. For example, during simulated medical diagnoses, the think-aloud protocols of senior medical students refer to biomedical knowledge more frequently than do the protocols of beginning medical students. However, seemingly paradoxically, the protocols of experienced physicians, like those of the beginners, contain little spontaneous mention of biomedical facts (e.g., Boshuizen & Schmidt, 1992; see also Schmidt & Boshuizen, 1993; van de Wiel, Boshuizen, & Schmidt, 2000). It turns out that this decreased mention of biomedical knowledge does not reflect its loss or inaccessibility, as revealed by the fact that the experts could be prompted at a later point to provide an extensive biomedical analysis of their diagnosis that far surpassed the detail offered by the medical students. This confirms that biomedical knowledge is essential to medical diagnosis at all levels of expertise, although its role becomes increasingly more tacit (Boshuizen & Schmidt, 1992). It follows that any claim that people discard old structures after acquiring new ones must be treated with some caution unless it is based on particularly extensive analysis.

On balance, then, we suggest that people typically retain old structures after learning new ones, in accord with the accretionist approach to development (e.g., Keil, 1994). As Johnson and Carey (1998) noted, this does not mean that radical restructuring cannot occur but simply that many forms of learning, including that seen in our experiments, do not result in the dissolution of prior knowledge.

The Role of Error

Representational change during learning occurs for a reason; our data suggest that one important reason is error. Without error, there is no restructuring. If error is present, then restructuring may occur but, according to our experiments, only when information about alternatives is available. Previous research in categorization (Ross, 1997) and categorical problem solving (Dixon & Bangert, 2002) reinforce this conclusion.

Ross (1997) described seven experiments in which he trained participants on a categorization task and then alerted them to additional information about the stimulus items. In some cases (Experiments 1–4), this information was related to category membership, whereas in other cases (Experiments 5–7), it was not. Participants demonstrably altered their responses to transfer stimuli on the basis of this additional information, even when it was irrelevant. Participants were performing quite poorly on the categorization task; their use of the new stimulus information in the search for alternative strategies to improve performance is therefore consistent with the error view.

Dixon and Bangert (2002) presented an extensive study in which third graders, sixth graders, and college students were compared in their ability to solve a series of gear problems. In a gear problem, participants must analyze a chain of interconnected gears to determine the direction of movement of a target gear (either clockwise, counterclockwise, or jammed). Several strategies can be brought to bear on this task. For example, akin to finger arithmetic, people can trace the sequence of directions with their fingers, people can categorize the gears using an alternating sequence (e.g., clockwise–counterclockwise–clockwise . . . .), they can skip every other gear during categorization (e.g., clockwise–skip–clockwise . . . .), or they can resort to an efficient strategy based on parity (odd numbers of gears maintain direction of rotation of the first gear in the sequence; even numbers reverse direction). Dixon and Bangert used think-aloud protocols to identify participants' strategy on each trial and then sought to predict strategy shifts from measures such as accuracy and latency on previous trials.

In line with the error account, Dixon and Bangert (2002) found that discovery of the simplest tracing strategy could be predicted on the basis of error: Regardless of age, people were more likely to discover the strategy when they took exceedingly long to respond to previous trials (all participants) or when they were inaccurate on those trials (only for participants who had previously used a more complex strategy). Moreover, this error-driven restructuring occurred only if people previously used some strategy, even if incorrect, as opposed to merely guessing the answer. This is in accord with our reveal manipulation, which provided externally what Dixon and Bangert's participants were able to produce for themselves.

Finally, we must note some limitations to the role of error. First, it is obvious that people's sensitivity to error is not fixed. Motivational factors, including payoffs for performance, perceived experimenter demands, individual differences, and so on certainly play a role in determining a participant's propensity to try to reduce errors. This fact is a problem for any theory of error-driven learning, of course, and is generally handled by constructs such as learning rates or error discounting. In the context of knowledge restructuring, it follows that whereas error is necessary for restructuring, the level of error that would be sufficient must vary on a case-by-case basis.

A potentially more serious limitation of our error account lies in the apparent propensity of people to restructure their knowledge even (or especially) when their current strategy seems to be error free. The study by Dixon and Bangert (2002), which clearly demonstrated the existence of error-driven knowledge restructuring, also uncovered a process they termed *redescription*. It turns out that redescription occurs not when performance is fraught with error but rather when it is highly accurate.

Beyond Error: Revelation and Redescription

We concluded earlier that whereas the revelation of additional information about a task was a necessary condition for a shift from a simple to a complex strategy, a shift from a rule-like strategy to a strategy based on instance memories may occur without revelation. The error account is limited in an analogous manner because there are occasions in which knowledge restructuring arises by redescription when performance is flawless.

Dixon and Dohn (2003) define redescription as "discovering and representing information embedded in one's current representation rather than from an external source" (p. 1083), which "produces representational change when one's current representation is highly accurate" (Dixon & Bangert, 2002, p. 919). In support, Dixon and Bangert noted that people shifted from the tracing strategy discussed above to the more efficient categorization strategy (clockwise–counterclockwise–clockwise . . . .) only when performance on the preceding trials was accurate and involved concentrated use of the tracing strategy. Dixon and Dohn (2003) further examined the redescription concept using a near-isomorph of the gear problem (viz., connected balancing beams) and likewise found that people switched from tracing to categorization even when the former supported near-perfect performance. Dixon and Dohn also found that redescription facilitated subsequent transfer of the categorization strategy to the gear problem, a finding that they attributed to redescription creating a more abstract problem representation. Overall, there is little doubt that error-free restructuring can occur when people repeatedly (and successfully) use a strategy whose problem representation contains information about an alternative way of approaching the task.

Why, then, did we not observe error-free restructuring in our experiments? Why did participants in the R100 and R90 conditions dismiss the additional information that would have enabled them to approach the task differently? The answer is implied by the central notion of redescription, namely, that restructuring involves the discovery of information that is embedded in the original representation of the task. In our studies, after people successfully used the label to classify items, there was no further information embedded in the label that people could have discovered. Accordingly, we did not observe redescription-based restructuring.

Knowledge Restructuring Versus Coexisting Alternatives

Our focus has been on knowledge restructuring, that is, the discovery of an alternative followed by a permanent shift to that new way of approaching the task. This issue is related to, but often considered different from, research on the continued coexistence of alternative strategies and people's ability to switch between those strategies (e.g., Shrager & Siegler, 1998). Bearing in mind this qualification, the link between knowledge restructuring and related investigations on the coexistence of strategies deserves to be explored.

In concept acquisition paradigms such as categorization and function learning, recent work on knowledge partitioning (e.g., Kalish, Lewandowsky, & Kruschke, 2004; Lewandowsky, Kalish, & Ngang, 2002; Lewandowsky & Kirsner, 2000; Yang & Lewandowsky, 2003, 2004) has shown that people sometimes prefer to acquire several partial expedient solutions rather than a single complex strategy. For example, Yang and Lewandowsky (2003, 2004; see also Lewandowsky et al., 2002) showed that people learn separate and independent solutions for components of a categorization task if each task component is associated with different context cues during training. Specifically, the category structure used by Yang and Lewandowsky resembled the one used in the present experiments, except that stimuli surrounding each linear component boundary were associated with a unique cue that constituted a third, normatively irrelevant, "context" dimension (akin to the label in our experiments, except that, here, the label probabilistically identified category membership, whereas in Yang and Lewandowsky's studies, context identified not a category but the partial boundary of the space). Yang and Lewandowsky showed that people learned each component boundary separately and applied it to any stimulus (even items located on the far side of the other component boundary) that was accompanied by the appropriate context cue. Paralleling the results of the present control conditions, participants in the studies by Yang and Lewandowsky (2003, 2004) continued to use their multiple expedient solutions throughout training without any evidence of subsequent integration.

Conclusions

We have presented three experiments that converge on a clear conclusion: People restructure their knowledge only when an initial expedient strategy gives rise to error and an alternative strategy is revealed that offers better performance. Each condition is necessary, but they are only jointly sufficient. When the expedient strategy works well, either because a single dimension has sufficient validity or because exceptions to a single-dimensional rule can be memorized, people resist restructuring despite experimental encouragement. Conversely, even if the expedient strategy gives rise to error, no restructuring occurs without revelation of information about the complex strategy. Further, we observed a dissociation between responses to complete stimuli on the one hand (where we can find restructuring) and partial stimuli on the other hand (where contingency judgments do not change).

References

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35,* 127–145.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates, and novices. *Cognitive Science, 16,* 153–184.

Bourne, L. E., Healy, A. F., Parker, J. T., & Rickard, T. C. (1999). The strategic basis of performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory and Language, 41,* 223–252.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General, 120,* 278–287.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology, 21,* 13–19.

Compton, B. J., & Logan, G. D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition, 19,* 151–158.

Davies, S. P. (1994). Knowledge restructuring and the acquisition of programming expertise. *International Journal of Human-Computer Studies, 40,* 703–726.

Demetriou, A., & Raftopoulos, A. (1999). Modeling the developing mind: From structure to change. *Developmental Review, 19,* 319–368.

Dixon, J. A., & Bangert, A. S. (2002). The prehistory of discovery: Precursors of representational change in solving gear problems. *Developmental Psychology, 38,* 918–993.

Dixon, J. A., & Dohn, M. C. (2003). Redescription disembeds relations: Evidence from relational transfer and use in problem solving. *Memory & Cognition, 31,* 1082–1093.

Galili, I., Bendall, S., & Goldberg, F. (1993). The effects of prior knowledge and instruction on understanding image formation. *Journal of Research in Science Teaching, 30,* 271–301.

Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology, 45,* 482–553.

Johnson, S. C., & Carey, S. (1998). Knowledge enrichment and conceptual change in folk biology: Evidence from Williams Syndrome. *Cognitive Psychology, 37,* 156–200.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111,* 1072–1099.

Keil, F. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). New York: Cambridge University Press.

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology, 51,* 53–89.

Lewandowsky, S., Kalish, M., & Griffiths, T. L. (2000). When several rules compete during categorization: Expedient errors and resistance to knowledge restructuring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1666–1684.

Lewandowsky, S., Kalish, M., & Ngang, S. K. (2002). Simplified learning in complex situations: Knowledge partitioning in function learning. *Journal of Experimental Psychology: General, 131,* 163–193.

Lewandowsky, S., & Kirsner, K. (2000). Expert knowledge is not always integrated: A case of cognitive partition. *Memory & Cognition, 28,* 295–305.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 37–50.

Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General, 117,* 68–85.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53–79.

Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education, 86*(4), 548–571.

Pearsall, N. R., Skipper, J. J., & Mintzes, J. J. (1997). Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. *Science Education, 81,* 193–215.

Posner, G., Strike, K., Hewson, P., & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66,* 211–227.

Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37,* 240–267.

Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General, 133,* 63–82.

Rumelhart, D., & Norman, D., (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–53). Hillsdale, NJ: Erlbaum.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory & Cognition, 21,* 338–351.

Shrager, J., & Siegler, R. S. (1998). A model of children's strategy choices and strategy discoveries. *Psychological Science, 9,* 405–410.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116,* 250–264.

van de Wiel, M. W. J., Boshuizen, H. P. A., & Schmidt, H. G. (2000). Knowledge restructuring in expertise development: Evidence from pathophysiological representations of clinical cases by students and physicians. *European Journal of Cognitive Psychology, 12,* 323–355.

Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research, 57,* 51–67.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11,* 381–419.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology, 43,* 337–375.

Yang, L.-X., & Lewandowsky, S. (2003). Context-gated knowledge partitioning in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 663–679.

Yang, L.-X., & Lewandowsky, S. (2004). Knowledge partitioning in categorization: Constraints on exemplar models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1045–1064.

Received August 6, 2004
Revision received November 16, 2004
Accepted February 23, 2005 ■