

Exemplar-based inference in multi-attribute decision making: Contingent, not automatic, strategy shifts?

Linnea Karlsson*

Department of Psychology
Umeå University

Peter Juslin

Department of Psychology
Uppsala University

Henrik Olsson

Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development

Abstract

Several studies propose that exemplar retrieval contributes to multi-attribute decisions. The authors have proposed a process theory enabling a priori predictions of what cognitive representations people use as input to their judgment process (*Sigma*, for “summation”; P. Juslin, L. Karlsson, & H. Olsson, 2008). According to *Sigma*, exemplar retrieval is a back-up system when the task does not allow for additive and linear abstraction and integration of cue-criterion knowledge (e.g., when the task is non-additive). An important question is to what extent such shifts occur spontaneously as part of automatic procedures, such as error-minimization with the Delta rule, or if they are controlled *strategy* shifts contingent on the ability to identify a sufficiently successful judgment strategy. In this article data are reviewed that demonstrate a shift between exemplar memory and cue abstraction, as well as data where the expected shift does *not* occur. In contrast to a common assumption of previous models, these results suggest a controlled and contingent strategy shift.

Keywords: exemplar memory, cue abstraction, strategy shifts, multi-attribute decisions, *Sigma*.

1 Introduction

Exemplar models have proven particularly successful for describing categorization learning and categorical decisions (Nosofsky & Johansen, 2000; cf. Medin & Schaffer, 1978). In the last years, researchers have shown interest in studying exemplar-based reasoning as a basis for other sorts of judgments and decisions (e.g., DeLosh, Busemeyer, & McDaniel, 1997; Dougherty, Gettys, & Ogden, 1999; Enkvist, Newell, Juslin, & Olsson, 2006; Juslin & Persson, 2002; Juslin, Olsson, & Olsson, 2003; Juslin, Jones, Olsson, & Winman, 2003; Juslin, Karlsson, & Olsson, 2008; Karlsson, Juslin, & Olsson, 2007; Nilsson, Olsson, & Juslin, 2005; Olsson, Enkvist, & Juslin, 2006; Olsson, Juslin, & Olsson, 2006; Sieck & Yates, 2001; Smith & Zarate, 1992).

It has thus been increasingly acknowledged that many

everyday judgments, like, for example, a diagnosis of a patient, or the decision to hire a job candidate or to purchase a car may, at least in part, be driven by the retrieval of concrete, similar previous examples. Despite this interest, the *nature* of the interaction between exemplar memory and other processes, such as rule-based processes, has received relatively little empirical scrutiny.

As an illustration, imagine a teacher predicting the final grades of students in an Economy class on the basis of four different test results. When looking at Robin, the teacher might infer that his high score on the Winter Exam should be associated with a high final grade. His medium score on the National High School Economy Exam should also be associated with a high grade, since that exam is very difficult. The teacher considers the different individual test results and integrates their combined impact on the grade. This exemplifies a judgment process based on abstracted knowledge of the relations between individual cues (the tests and their difficulty) and the criterion (the final grade).

For Lisa, on the other hand, the teacher might realize that her test results are very similar to Robin's. Since the teacher gave Robin a final grade of 8 out of 10, she de-

*Linnea Karlsson is now at the Max Planck Institute for Human Development, Berlin. This research was supported by Bank of Sweden Tercentenary Foundation and the Swedish research council. Correspondence concerning this paper should be addressed to: Linnea Karlsson, Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Lentzeallee 94, 141 95 Berlin, Germany, or to the e-mail address: karlsson@mpib-berlin.mpg.de.

cides to give Lisa an 8 as well. This illustrates a judgment based on a similar concrete exemplar. Although these examples are highly simplified, we can probably all agree that we have at times engaged in both sorts of processes. What is less clear, and has received relatively little attention in judgment and decision research, are what task properties invite the one or the other process, and what processes that instigate shifts between the processes.

The possibility that exemplar memory interacts with rule-based processes has been modeled in the context of, for example, classification learning (e.g. Bourne et al., 1999; Erickson & Kruschke, 1998; Palmeri, 1997), and arithmetic learning (Rickard, 1997, 2004; Logan, 1988). Without claiming identity, these models share the assumption that a transition from one process to another occurs automatically at an item-based level, either as a side-effect of obligatory, accumulative encoding (Logan, 1988) or as implemented by an error-driven learning rule typical of connectionist models (e.g., Erickson & Kruschke, 1998).

A viable alternative, however, is that the shifts are hypothesis-driven, involving controlled and sequential tests of strategies, sustained over a period of trials (Brehmer, 1994; Haider, Frensch, & Jarom, 2005; Meeter, Myers, Shohamy, Hopkins, & Gluck, 2006). According to this view, the choice of the one rather than the other strategy is the result of an active and effortful process of problem solving and contingent on the perceived success of the strategy.

The aim of this article is to review data on *multiple-cue judgment* (Cooksey, 1996; Hammond & Stewart, 2001) that demonstrate abilities to shift between the processes, as well as striking and stubborn inability to make appropriate shifts. These data, in turn, serve to shed light on the nature of the mechanisms controlling the shifts between processes.

2 Exemplar memory

That experience is partly structured in the form of episodic memories is an acknowledged part of cognitive psychology (Tulving, 1972). You can recall a certain event that has occurred to you and you can state that you have met a certain brilliant scientist once before. The concept of *similarity* is likewise a classic and important phenomenon in cognitive psychology (Shepard, 1987). You may have a belief that your grandmother's painting is similar to one you have seen at Guggenheim and that your new patient makes similar complaints as the patient you met yesterday. The concepts of episodes and similarity unite in the exemplar-based framework (Estes, 1994; Medin & Schaffer, 1978; Nosofsky & Johansen, 2000). An exemplar-based process involves a comparison be-

tween a judgment probe and similar instances stored in memory. Exemplar-based reasoning allows flexible inferences in the sense that minimal commitments to the underlying task structure is made at the original time of encoding (Juslin & Persson, 2002, see Aha, 1997 on "Lazy Algorithms").

Medin and Schaffer (1978) provided one of the most influential mathematical formulations of exemplar theory; the *context model* (see also Nosofsky, 1984, 1986, for the further development into the Generalized Context Model). With the context model it is assumed that when a judgment is made about a probe, the judge considers the similarity of the probe to all or some of the previously encountered exemplars. Similarity acts as a weight on the stored criterion values. When applied to a continuous criterion in a multiple-cue judgment task (e.g., Juslin, Olsson et al., 2003), a similar exemplar suggests a similar criterion value, whereas a dissimilar receives less weight in the judgment. The weighted criterion values are then added and divided by the sum similarity of all exemplars considered. The similarity between a probe and an exemplar is determined by feature overlap. An exemplar with large feature overlap receives a relatively dominant impact on the judgment (for details, see below).

As already noted, exemplar memory has proven successful in accounting for a wide range of judgment phenomena, including *categorization* (e.g., Nosofsky & Johansen, 2000), *function learning* (DeLosh et al., 1997), *implicit learning* (Pothos & Bailey, 2000), *likelihood judgment* (Dougherty et al., 1999), *multiple-cue judgment* (Enkvist et al., 2006; Juslin, Olsson et al., 2003; Juslin, Jones et al., 2003; Juslin et al., 2008; Karlsson et al., 2007; Olsson, Enkvist et al., 2006; Olsson, Juslin et al., 2006), *confidence* (Juslin & Persson, 2002; Nilsson et al., 2005; Sieck & Yates, 2001) and *social judgment* (Smith & Zarate, 1992).

3 Exemplar memory and its interplay with other processes

Contemporary conceptions of the interplay between exemplar memory and rule-based or algorithm-based processes often frame the shift between processes as an "inevitable consequence of task experience" (as characterized by Haider et al., 2005, p. 496; see e.g. Bourne et al., 1999; Erickson & Kruschke, 1998; Logan, 1988; Palmeri, 1997; Rickard, 1997, 2004). This suggests that the shifts arise as parallel, more or less, automatic (cf. Shiffrin & Schneider, 1977) side-effects of judgment practice.

3.1 Obligatory encoding and automatic competitive retrieval

One proposal is that rule-based processes compete with the retrieval of individual instances, where further training leads to gradual accumulation of individual instances that ultimately come to dominate the response output. In the theory of automatization, Logan (1988) proposes that the key to skilled performance is memory:

“The theory makes three main assumptions: First, it assumes that encoding into memory is an obligatory, unavoidable consequence of attention. Attending to a stimulus is sufficient to commit it to memory. It may be remembered well or poorly, depending on the conditions of attention, but it will be encoded. Second, the theory assumes that retrieval from memory is an obligatory, unavoidable consequence of attention. Attending to a stimulus is sufficient to retrieve from memory whatever has been associated with it in the past. Retrieval may not always be successful, but it occurs nevertheless. Encoding and retrieval are linked through attention; the same act of attention that causes encoding also causes retrieval. Third, the theory assumes that each encounter with a stimulus is encoded, stored, and retrieved separately.” (p. 493).

This implies that as experience with a task increases there will be more and more available instances stored. Logan (1988) proposed that the transition from what he called “algorithm-based processing” to instance retrieval occurs as a competition and the process that can produce the fastest output “wins”. Rickard (1997) makes similar proposals of a competition between algorithm-based and instance-based processes that race to produce the output in every trial. He modified Logan’s account somewhat and assumed that in every trial one or the other process will be strengthened, and the process that is most stable at the moment will be responsible for the response. In a model called *EBRW* (Exemplar-Based Random Walk) Nosofsky & Palmeri (1997) proposed that learning in a perceptual judgment task proceed as a gradual accumulation of exemplars as suggested by Logan.

3.2 Procedures for error minimization

Another influential idea is that there is an adaptive mechanism that, after each consecutive judgment, adapts the weight given to different processes in a manner that should minimize the judgment error on the next judgment trial, much in the spirit of conditioning and reinforcement learning. This mechanism often involves some version of the Delta-rule, or its derivative Back-propagation (Ellis & Humphreys, 1999). In categorization learning, a common assumption has thus been that category learning is an error-driven competition between processes (e.g., Ashby

et al., 1998; Erickson & Kruschke, 1998; Palmeri, 1997).

Erickson and Kruschke (1998) propose a sophisticated connectionist model (ATRIUM) constructed as two separate modules; one *rule-module* and one *exemplar-module*. Both modules process the incoming stimuli, in parallel, but the module with the highest activation will give rise to the output. After each judgment, the activation of the modules is adjusted to allow better future performance. Ashby et al. (1998) likewise suggests a competitive interplay between an explicit verbal system and an *implicit procedural system* (and thus not exemplar-based in the episodic sense) that is driven by the rate of categorization error.

3.3 Controlled and contingent strategy shifts

Another possibility is that learning to perform well in a judgment task is primarily governed by explicit and controlled attempts on behalf of the learner. Haider and colleagues (e.g., 2005) initiated a series of experiments aiming at identifying voluntary components to strategy shifts. The task paradigm they used — the Alphabet Verification Task (AVT) — demands a judgment of whether strings of letters are alphabetically correct or not. A typical string could look like CDEFG[4]L, where the digit in brackets is to be interpreted as “is there four letters between G and L” (this rule was not told to the participants beforehand). Participants started with a strategy where they focused on all elements in the string but shifted to a strategy where they only considered the triplets involving the digit (i.e. G[4]L).

Haider and colleagues interpreted the data as support for an abrupt adoption of the later strategy, not a gradual transition from one to the other. They suggest that a factor inducing the abrupt shift was that the participants became aware that during some trials the responses were faster than during others: “The abruptly occurring violations of expectation, we presently assume, might serve as triggers for explicit inferential processes” (2005, p. 517). Recent data from probabilistic category learning likewise appear to provide additional empirical support for voluntary shift components (Meeter et al., 2006). These authors assume that participants do discrete switches between different strategies as they try to solve the task. The individual data are modelled in order to try to identify the hypothesized “switch points”. Similar discrete switching results have been reported by Rehder & Hoffman (2005).

In regard to multiple-cue judgment, a possibility is that different strategies for performing the judgments are tried one at the time, much as in a *lexicographic order*, until a strategy is found that yields satisfying performance. A candidate strategy is thus pursued to the extent that it is perceived by the judge to deliver acceptable judg-

ment performance; if the initial performance is too poor, another strategy is selected. While one strategy is attempted, little or no learning with respect to the other strategies occurs (e.g., as long as the judge pursues the strategy of cue abstraction, little information relevant to the alternative strategy of exemplar memory is accumulated). While little is known about such a hypothetical lexicographic ordering of processes or strategies, the common observation of a preference for explicit rule-based processes (a “rule bias”, cf. Ashby et al., 1998) suggest that rule-based strategies of cue abstraction are attempted before exemplar memory.

There are, at least, two differences between this hypothesis of a controlled and contingent mechanism and the previous more automatic mechanisms for a shift between the processes. First: as noted, the knowledge gained is primarily related to the strategy that is actively pursued. Second: the continued application of a strategy is contingent on its ability to deliver acceptable performance early in learning; otherwise the judge will shift to reliance on some other strategy. Together these properties suggest that learning is not an inevitable consequence of experience, but the outcome of *active problem solving* by the judge.

Berndt Brehmer formulated a well-known similar “hypothesis testing” model in the context of learning the functional relations between continuous cues and criteria; participants evaluate different hypotheses regarding the function relating the cues to the criteria in a lexicographic order. A positive, linear function relating the cues to a criterion has been argued to be the first hypothesis that is tested by a function learner (Brehmer, 1994). The present proposal extends such a cognitive approach to the choice of judgment strategy.

Next, we will describe the judgment task used in the studies we review, followed by an overview of Sigma, the framework in terms of which we derive and test predictions concerning representational shifts in multiple-cue judgment. Thereafter, we will review data on peoples’ abilities to shift between representations in variations of this task and discuss the implications for a putative mechanism underlying adaptive shifts between different cognitive strategies.

4 A multiple-cue judgment task

In order to distinguish between exemplar memory and cue abstraction we have employed a specific task design (see e.g., Juslin, Olsson et al., 2003). The idea is to provide a learning task in which some exemplars are withheld, only later to be introduced in a test phase. In this way, we can study the judgment performance for the new exemplars which depends on the cognitive processes used

to make the judgments.

All the experiments reviewed in this article involve judgments of the toxicity of subspecies of a fictitious poisonous Death Bug (in the case modifications are made to the task, they are described in connection with the experiment in question). The bugs come in different species that vary on four binary cues (e.g., length of the legs, pattern on the back). The cues either have a value of 0, implying no change on the criterion, or a value of 1 implying a change (see Table 1). In one version of the task the bugs are displayed as drawings, and in one version the cue values are presented as four written propositions.

The task structure is summarized in Table 1. In a binary version of the task, half of the bugs are assigned to a harmless category (criterion = 0) and the other half are assigned to a dangerous category (criterion = 1). In an additive continuous version of the task, the cue-combination rule is strictly a weighted linear function of the four cues,

$$c = 50 + 4 \cdot C_1 + 3 \cdot C_2 + 2 \cdot C_3 + 1 \cdot C_4 \quad (1)$$

C_1 is the most important cue with a *coefficient* of 4 (i.e., a relative weight .4), C_2 is the second most important cue with a coefficient of 3, and so forth. In Table 1, a subspecies with feature vector [0, 0, 0, 0] therefore has a toxicity of 50%; a subspecies with feature vector [1, 1, 1, 1] has 60%. The other variations of the task displayed in Table 1 will be further explained below. For some of the experiments reported below a normally distributed random error was added to the error-free criterion, transforming the discrete 11 values in Table 1 into a truly continuous and *probabilistic* variable (correlation $r = .9$ with the error-free criterion).

In a learning phase, the participants learn to make judgments from outcome feedback. The learning phase consists of 11 of the subspecies in Table 1 (denoted “O” in Table 1). After the learning phase there is a test phase where all the 16 exemplars from Table 1 are included. Judgments on the 5 new exemplars (denoted “E” and “N” in Table 1) are important for identifying what cognitive process that might have been at play.

5 Sigma: predictions for cue abstraction and exemplar memory

The present authors have proposed a cognitive theory of multiple-cue judgment, predicting a sophisticated division of labor between exemplar memory and cue abstraction, in a model called *Sigma* (Juslin et al., 2008). Juslin et al. (2008) suggest that because of the constraints on controlled judgment processes, the abstraction and integration of the separate impact of distinct cues on the criterion is limited to tasks where the cues combine linearly and additively, and to tasks where such a model is

Table 1: The 16 Exemplars with their Cues and Criteria Prior to Addition of Random Error for the tasks in Experiment 1, 2 and 3 (Juslin, Olsson et al., 2003), from Experiments 1 & 2 (Juslin et al., 2008) and Experiments 1 and 2 (Olsson, Enkvist et al., 2006).

Exemplar #	Cues				Criterion			Role	
	C ₁	C ₂	C ₃	C ₄	Binary	Add	Mult		Non-linear
1	1	1	1	1	1	60	72.75	50	E
2	1	1	1	0	1	59	59.00	53.6	T
3	1	1	0	1	1	58	53.94	56.4	T
4	1	1	0	0	1	57	52.08	58.4	O
5	1	0	1	1	1	57	52.08	58.4	N
6	1	0	1	0	1	56	51.40	59.6	N
7	1	0	0	1	$p = .5$	55	51.15	60	N
8	1	0	0	0	0	54	51.15	59.6	T
9	0	1	1	1	1	56	51.40	59.6	O
10	0	1	1	0	$p = .5$	55	51.14	60	O
11	0	1	0	1	0	54	51.05	59.6	T
12	0	1	0	0	0	53	51.02	58.4	T
13	0	0	1	1	0	53	51.02	58.4	T
14	0	0	1	0	0	52	51.00	56.4	T
15	0	0	0	1	0	51	51.00	53.6	T
16	0	0	0	0	0	50	51.00	50	E

a good approximation. In other tasks, exemplar memory is suggested to act as a back-up system that is more independent of the specific task structure (e.g., the manner in which the cues combine). This knowledge, however, comes at the cost; the task structure is not explicitly analyzed and ready for verbalization.

For a complete and detailed proposal of Sigma readers are referred to Juslin et al. (2008). There are three main assumptions of the theory:

Assumption 1. The judgment process is inherently constrained to successive consideration of two real or potential estimates of the criterion. At the time of learning, this constraint implies that the problem solving involved in estimating the weight of individual cues (i.e., cue abstraction) is restricted to the comparison of two exemplars with different criteria and different cues. For example, observing exemplars [0, 1, 1, 1] with criterion 56 in close sequel to exemplar [1, 1, 1, 1] with criterion 60 may suggest that the first cue has the effect of adding four units to the criterion.

This constraint thereby imposes a severe difficulty to infer non-linear cue-criterion relations, since for such inferences at least three data points are needed (e.g., at least

three data-points are needed in a bi-variate plot in order to discover a nonlinear function).

At the time of judgment, the constraint to only consider two estimates implies a judgment process where the estimate at one time is adjusted into a new estimate on the basis of the consideration of additional evidence (i.e., a cue or a similar exemplar). This constraint thus implies a sequential adjustment of a tentative estimate that implements an *additive* combination of the impact of each cue (i.e., each cycle involves the adjustment of a previous estimate \hat{c}_{n-1} into a new estimate \hat{c}_n).

Assumption 2. The judgment process may be driven by different representations, for example, either by abstract knowledge of cue-criterion relations or by memory for concrete previous exemplars.

Assumption 3. The selection of input to the process is not arbitrary, but tends to shift to an input appropriate to the task at hand. The third assumption is accordingly the assumption scrutinized in this article. Figure 1 illustrates the sequential judgment process, as hypothesized in Sigma.

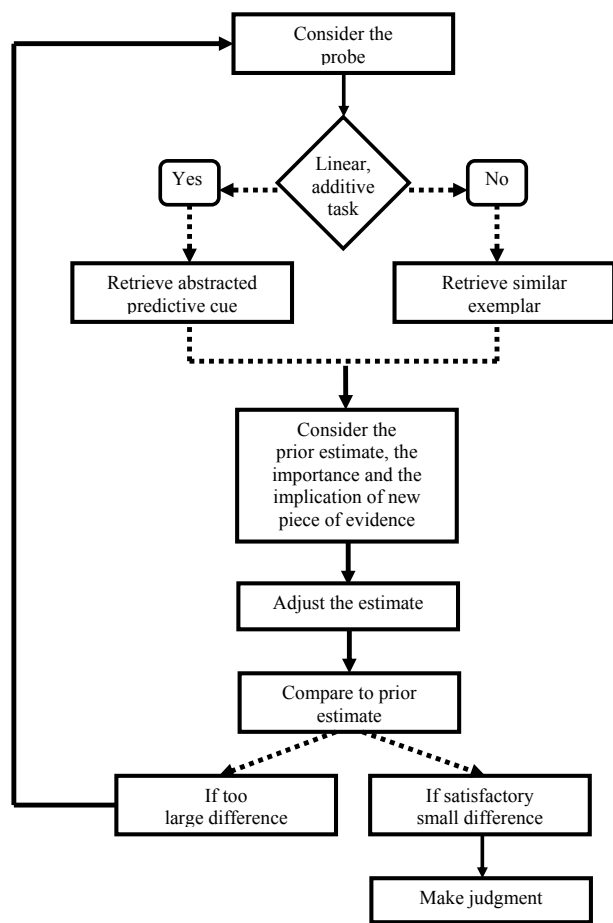


Figure 1: A flow-chart exemplifying the iterative and sequential judgment cycle hypothesized by Sigma (for the complete mathematical formulation, see Juslin et al., 2008). When making a judgment about a probe, the information input to the process at a certain time point is either an abstracted predictive cue or a similar exemplar, depending on the structure of the task. An adjustment of a prior estimate of the probe is made (if such exists, otherwise of a default estimate), by considering the importance of the retrieved piece of evidence, as well as the estimate it implies. This procedure of sampling pieces of evidence from memory continues until the judge perceives the estimate to be good enough, and a judgment can be made.

By considering cognitive limitations in relation to the task we can predict *task-dependent* shifts in the process. Sigma adheres to the growing body of evidence indicating that controlled cognitive processing is constrained to additive and serial consideration of information (Anderson, 1981; Chapman & Johnson, 2002; Cooksey, 1996; Denrell, 2005; Fischbein, Deri, Nello, & Marino, 1985; Hammond & Stewart, 2001; Hogarth & Einhorn, 1992; Roussel, Fayol, & Barrouillet, 2002; Shiffrin & Schnei-

der, 1977).

This implies that human judgment is *bounded* (Simon, 1990), not only in the time and knowledge available, or with regard to computational ability in some unspecified sense, but with regard to the *information integration* that can be performed (see Anderson, 1981). Sigma thus emphasizes that the abstraction and integration of cues in multiple-cue judgment must conform to this constraint. The task structure is crucial; to abstract and integrate the additive impact of cues on a criterion will not produce accurate judgment in tasks where the cues combine in a highly non-additive or non-linear manner. Because we are able to learn also in such tasks, we have to resort to exemplar memory instead (Juslin et al., 2008). Interestingly, the claim that additive and non-additive tasks induce different cognitive processes has recently been implied also by brain imaging results (Karlsson, Nyberg, Juslin & Olsson, 2007).

The framework of Sigma can also be used to predict *time-dependent* shifts. On the assumption that there exists a rule-bias (e.g., Ashby et al., 1998) the participants will start with cue abstraction as they approach a task. If the task is well approximated by linear, additive integration of the cues, cue abstraction will continue to be a viable alternative. If not, there is a shift to exemplar memory. Sigma therefore serves as a framework for understanding the role of different cognitive processes in multiple-cue judgment. As with many models of category learning (e.g., Ashby et al., 1998; Erickson & Kruschke, 1998) it assumes that people shift adaptively between processes, but Sigma emphasizes that both cue abstraction and exemplar memory involve controlled cognitive processing components.

We now consider the quantitative predictions that can be derived from Sigma for the judgment tasks in Table 1, when it is fed either with representations of cue-criterion relations (cue abstraction) or with exemplars (exemplar memory). (Again, the reader is referred to Juslin et al., 2008, for details about the derivation of the predictions). These quantitative models allow us to test the key prediction by Sigma: in tasks where the criterion is not well predicted by a linear, additive model, we should observe a shift from cue abstraction to exemplar memory.

5.1 Cue abstraction

The estimate of the criterion by cue abstraction is equivalent to a linear additive model,

$$\hat{c} = a + b_I \cdot C_I + \dots + b_I \cdot C_I \tag{2}$$

where $a = 50 + .5 \cdot (10 - \sum b_i)$, I is the total number of cues, and the linear weights $b_i (i = 1 \dots i)$ are estimated by regression analysis (see Juslin et al., 2008, for further details, see also Juslin, Olsson et al., 2003). In the binary

criterion case, the estimate of the criterion by the cue abstraction model is equivalent to logistic regression,

$$p(b = 1) = \frac{e^{-.5 \sum_I w_i + \sum_I w_i c_i}}{1 + e^{-.5 \sum_I w_i + \sum_I w_i c_i}} \quad (3)$$

where $p(b = 1)$ is the probability of responding 1 (in our case “dangerous”) and w_i are logistic cue weights. We assume that crossover between binary judgment 0 to 1 occurs at toxicity 55, hence the term $-.5 \sum_I w_i$. In this task, the binary cue abstraction model is formally identical to a prototype model with a multiplicative similarity rule (Olsson & Poom, 2005).

Because the process is the same, cue-abstraction predicts that judgments for new exemplars will not deviate systematically from judgments for the old exemplars, and implies the ability to extrapolate beyond the training range of the criterion, even if these extreme exemplars have never been encountered in training (see Figure 2 A, C).

5.2 Exemplar memory

The estimate of the criterion by exemplar memory is a weighted average of the criteria of the retrieved exemplars (Juslin, Olsson et al., 2003),

$$\hat{c} = \frac{\sum_N S_n \cdot c_n}{\sum_N S_n} \quad (4)$$

where N refers to the number of exemplars, S_n refers to the probe-exemplar similarity and c_n to the criterion of exemplar n ($n = 1 \dots N$). The similarity between the probe and exemplar x_n is computed with the similarity rule of the original context model (Medin & Schaffer, 1978):

$$S_n = \prod_{i=1:I} d_i \quad (5)$$

d_i is an index that takes value 1 if the cue values on cue dimension i ($i = 1 \dots i$) coincide (i.e., both are 0 or both are 1), and s_i if they deviate (i.e., one is 0, the other is 1). s_i are four parameters in the interval [0, 1] that capture the impact of deviating cues on the similarity S_n . Equations 4 and 5 can be used both when the criterion is continuous and when it is binary.

With exemplar memory the prediction is that judgments for new exemplars will be poorer than for old exemplars, because old exemplars can benefit from retrieval of identical previous exemplars with the correct criterion. There will also be an inability to extrapolate outside of the training criterion range for the new extreme exemplars, because the judgment is a weighted average of the criterion values in training (Figure 2 B, D).

In sum, one account of multiple cue judgment is *cue abstraction*; the judge attends to individual cues in sequence, each with a known relationship to the criterion,

and mentally integrates them into an overall judgment (Einhorn, Kleinmutz, & Kleinmutz, 1979).¹ Another account is *exemplar memory*, where the judge retrieves similar exemplars from memory and integrates the criterion values that are stored together with the exemplars (Medin & Schaffer, 1978; Nosofsky & Johansen, 2000). Note that with exemplar-memory there has been no abstraction on the level of the individual cues and their relation to the criterion. Rather, it is the pattern of cues in relation to the criterion that is driving the judgments. Sigma predicts that people shift between cue abstraction and exemplar memory depending on the task.

As we shall see, however, there are limitations to this account. Recent results suggest that, despite the assumed flexibility of exemplar memory in relation to the task structure, participants do not always shift to exemplar memory in the predicted manner in tasks where cue abstraction is not a viable alternative. The scope of this article is to explore the theoretical implications of those results. We will argue that there are limitations to an “inevitable consequence of task experience” (Haider et al., 2005) account of the interplay and suggest that it is better described as a controlled *strategy shift* contingent on early learning performance.

6 The empirical studies

Sigma predicts that the task structure is one factor shaping the cognitive process that dominates in a judgment task. Four types of tasks were used to derive predictions that were tested in the experiments reviewed below: 1) a *binary task*, where the feedback is assumed to be too poor to encourage abstraction of the underlying cue-criterion relations, and which hence should invite exemplar memory (see Juslin, Olsson et al., 2003); 2) an *additive task* with a continuous criterion, which should allow, and invite, the abstraction and integration of multiple cues in a linear and additive manner (Juslin, Olsson et al., 2003; Juslin et al., 2008; Olsson, Enkvist et al., 2006); 3) a *multiplicative task*, where the cue-criterion relations are not well predicted by a linear, additive model and the process therefore should be exemplar memory (Juslin et al., 2008), 4) a *non-linear task*, which similarly should not invite additive, linear integration of cues, and hence should induce use of exemplar memory (Olsson, Enkvist et al., 2006).

¹Note that cue abstraction does not necessarily involve integration of cues. There might be situations where considering the individual impact of one single cue on the criterion is enough to perform well.

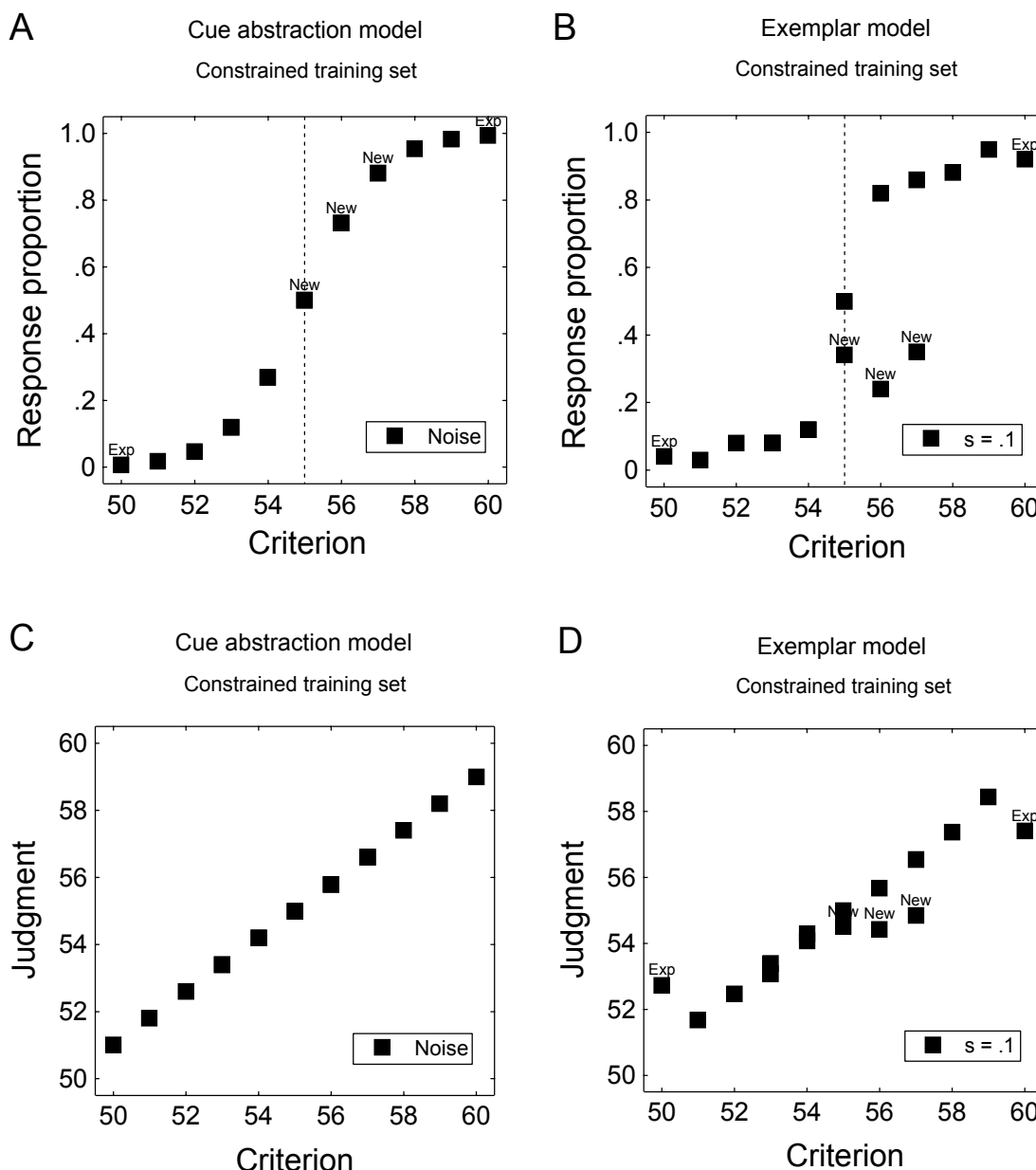


Figure 2: Predicted judgments for a binary (A, B) and a continuous (C, D) criterion with the constrained training set. Panel A and C: Cue abstraction model with noise for the constrained training set. Panel B and D: Exemplar model with similarity parameter $s = .1$.

6.1 General Method

Participants in all experiments reviewed were undergraduate students at Umeå University, Sweden. Stimuli and procedures were basically the same for all the experiments. The participants learned to make judgments of the toxicity of fictitious bugs, varying on four binary dimensions (i.e., length of their legs, spots or no spots on their fore back). The experiments consisted of a learning phase and a test phase. In the learning phase, participants

judged the toxicity of each of the bugs and received outcome feedback. The toxicity was either binary (“harmless” or “dangerous”; Experiment 1, Juslin, Olsson et al., 2003) or continuous varying from 51–59% toxicity in the learning phase (the other experiments in this review). The learning phase included 11 of the bugs from the stimulus set (Table 1) shown 20 times each. In the test phase, the task was again to judge the toxicity of each bug, but without outcome feedback. The test phase included all bugs from the stimulus set (16 bugs).

We will discuss three dependent measures: *performance*, *exemplar index* and *model fit*. Performance was measured in terms of root mean squared error (RMSE) between judgment and criterion. The exemplar index measures to what extent exemplar memory has dominated the judgments, and is a combination of an *extrapolation index* and an *interpolation index*. The deviation between the judgments for the new extreme exemplars introduced at test (i.e. those with cue values [0,0,0,0] and [1,1,1,1]) and the predicted linear extrapolation for those items, based on the old exemplars at test defines the *extrapolation index*. The *interpolation index* is the difference between absolute deviations between judgment and criterion for the new exemplars in the middle of the criterion range (i.e. exemplar nr 5, 6, & 7) and the matching old exemplars. The sum of the extra- and interpolation indices is the exemplar index. A negative exemplar index is indicative of exemplar memory, because it captures systematic deviations from the correct criterion value for the new exemplars, as predicted by the exemplar model. An exemplar index of 0 is indicative of cue abstraction, since this suggests no systematic difference in judgment error between the new and the old exemplars.

The models described above were used to predict the judgments from the test phase. The parameters were estimated for the judgments in the last part of the learning phase and used to predict data in the test phase (a method we call *projective fit*). The model fit is reported in terms of the *coefficient of determination* (r^2) and the *root mean square deviation* (RMSE) between the predictions and the data. The reader is referred to the original publications for additional details on the methods of the experiments.

7 Successful representational shifts

7.1 Binary vs. continuous criterion (Juslin, Olsson et al., 2003)

The first two experiments involve variations of the linear, additive version of the task (Eq. 1), with either a binary criterion or a continuous criterion (Experiments 1 & 2 in Juslin, Olsson et al., 2003). According to Sigma, the feedback is crucial for the ability to abstract explicit knowledge of cue-criterion relations. If one assumes a linear additive model (as people often do, e.g., Brehmer, 1994) and if the additive task described by Eq. 1 is deterministic, in principle observation of only five exemplars is sufficient to uniquely determine the five coefficients. By contrast, when the feedback is binary or nominal, often the relevant structure cannot be identified from a few observations, even if the correct model is assumed (Juslin, Olsson et al., 2003). Therefore, we predicted reliance on exemplar memory in the binary version of the

additive task and a shift towards cue abstraction when the criterion is continuous.

Experiment 1 involved judgments of a deterministic binary criterion and Experiment 2 judgments of a deterministic continuous criterion. Figure 3 shows the mean judgments plotted against the criterion (log odds transformed for the judgments in Experiment 1). Clearly, there are larger differences between new and old exemplars when the criterion is binary than when it is continuous. The proportion of participants for which the 95% confidence interval for the exemplar index includes 0, as predicted by cue abstraction (i.e., the judgments for new and old exemplars do not deviate in a systematic manner), was 28% with the binary criterion and 50 % with the continuous criterion. The shift towards more cue abstraction was also verified by the model fit indices. The fit was better for the exemplar model when the criterion was binary ($r^2 = .93$ with $RMSE = 0.089$ vs. $r^2 = .80$ with $RMSE = 0.18$) with a tendency toward the opposite pattern when the criterion was continuous ($r^2 = .89$ with $RMSE=0.63$ vs. $r^2 = .92$ with $RMSE = 0.58$). These results suggest that when the criterion was changed from a binary into a continuous criterion the processes shifted from exemplar memory to cue abstraction.

7.2 Deterministic vs. probabilistic criterion (Juslin, Olsson et al., 2003)

The third experiment in Juslin, Olsson et al. (2003) investigated the effect of a probabilistic relation between the cues and the criterion. A deterministic task allows for perfect accuracy by retrieving exemplars that have been observed previously and the reoccurring presentation of a few unique exemplars (i.e., the same cues and criterion) may promote exemplar memory. In a probabilistic task, on the other hand, the same cues and criteria do not reoccur, and exemplar memory does not allow perfect accuracy. This suggests that exemplar memory may become less prevalent in a probabilistic task and that people shift to cue abstraction.

Experiment 3 involved the same stimuli and task structure as Experiment 2 in Juslin, Olsson et al. (2003), but contrasted a deterministic condition with a probabilistic condition where the multiple correlation between the cues and the criterion was .9. The proportion of participants for which the 95% confidence interval for the exemplar index includes 0 was 44% with the deterministic criterion and 84 % with the probabilistic criterion. The shift towards more cue-abstraction was also verified by the model fit indices. The fit for the exemplar model was almost as good as the cue abstraction model in the deterministic condition ($r^2 = .90$ with $RMSE = 0.57$ vs. $r^2 = .92$ with $RMSE = 0.62$). While the exemplar model maintained its level of fit, the cue abstraction

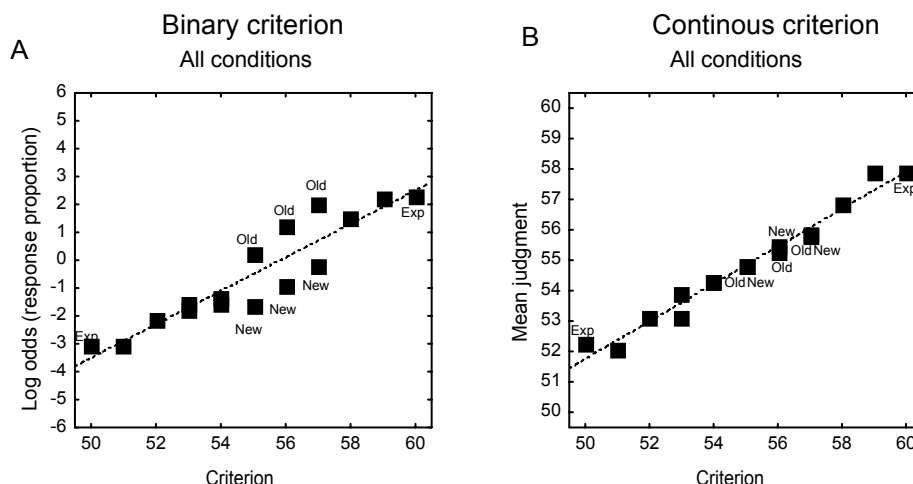


Figure 3: Mean judgments in a task with binary criterion (log odds transformed, Experiment 1) and a task with continuous deterministic criterion (Experiment 2). Adapted from “Exemplar Effects in Categorization and Multiple-Cue Judgment” by P. Juslin, H. Olsson, & A-C., Olsson, 2003. *Journal of Experimental Psychology: General*. Copyright by the American Psychological Association.

model became better at accounting for the data in the probabilistic condition ($r^2 = .90$ with $RMSD = 0.56$ vs. $r^2 = .95$ with $RMSD = 0.42$). The results of Experiment 3 suggest that cue abstraction increases in a task where the criterion is probabilistic.

More generally, the shift in processing observed across all experiments in Juslin, Olsson et al. (2003) indicates that changing a binary task into an additive task with a continuous criterion induces adaptive shifts from exemplar memory to cue abstraction. This is consistent with the success of exemplar models in categorization studies (e.g., Nosofsky & Johansen, 2000) and the assumption that multiple-cue judgment often involves cue abstraction (Einhorn et al., 1979), but also with a bias to abstract rules whenever possible (Ashby et al., 1998).

7.3 Additive vs. multiplicative cue-combination rules (Juslin et al., 2008)

Experiments 1 and 2 in Juslin et al. (2008) were designed to address a shift depending on whether the cue-combination rule is additive or multiplicative. As predicted by Sigma, cue-abstraction is a viable alternative when the task at hand is additive and linear. On the other hand, if the cues combine non-additively the abstraction of linear slopes between the cues and the criteria is not possible and, hence, we predict a dominance of exemplar memory in a distinctly non-additive task. The two experiments compared an additive version of the task (Eq. 1) with a multiplicative version of the task. In the multiplicative version the cue-combination rule was a non-

additive combination of the four cues,

$$c = 51 + 0.0009875 \cdot e^{4C_1+3C_2+2C_3+C_4} = 51 + 0.0009875 \cdot e^{4C_1} \cdot e^{3C_2} \cdot e^{2C_3} \cdot e^{C_4} \quad (6)$$

with the same coefficients as in the additive task (Eq. 1). A cue with value 0 leaves the expression in Eq. 6 unchanged, while a cue with value 1 multiplies the rest of the expression with a constant that is specific to each cue. For example, the multiplicative criterion for subspecies [1, 1, 1, 0] is $51 + 0.0009875 \cdot (54.60 \cdot 20.09 \cdot 7.39 \cdot 1) \approx 59.00$, and for subspecies [0, 0, 0, 1] it is $51 + 0.0009875 \cdot (1 \cdot 1 \cdot 1 \cdot 2.72) \approx 51.00$ (see Table 1).

The idea with Experiment 1 was to investigate whether we could observe a task dependent shift in the cognitive processes and the hypothesis was that we would observe more reliance on exemplar memory in the multiplicative task.² The judgments plotted against the criterion values are shown in Figure 4. There was better performance in the additive than in the multiplicative task. There was also a significant interaction between performance for the old as compared to the new exemplars in the additive and in the multiplicative tasks: the judgments for the old exemplars were about equally good in both tasks, while the judgments for the new exemplars were much worse in the multiplicative task (Figure 4C).

²The original experiment involved a dual-task manipulation of working memory and therefore had two test-phases. The secondary task involved listening to Swedish nouns read from loudspeakers and to count those nouns denoting something “alive”. Since we are only interested in the main effect of task structure for this review, the reader is referred to the original manuscript for further discussion of the results (Juslin et al., 2008).

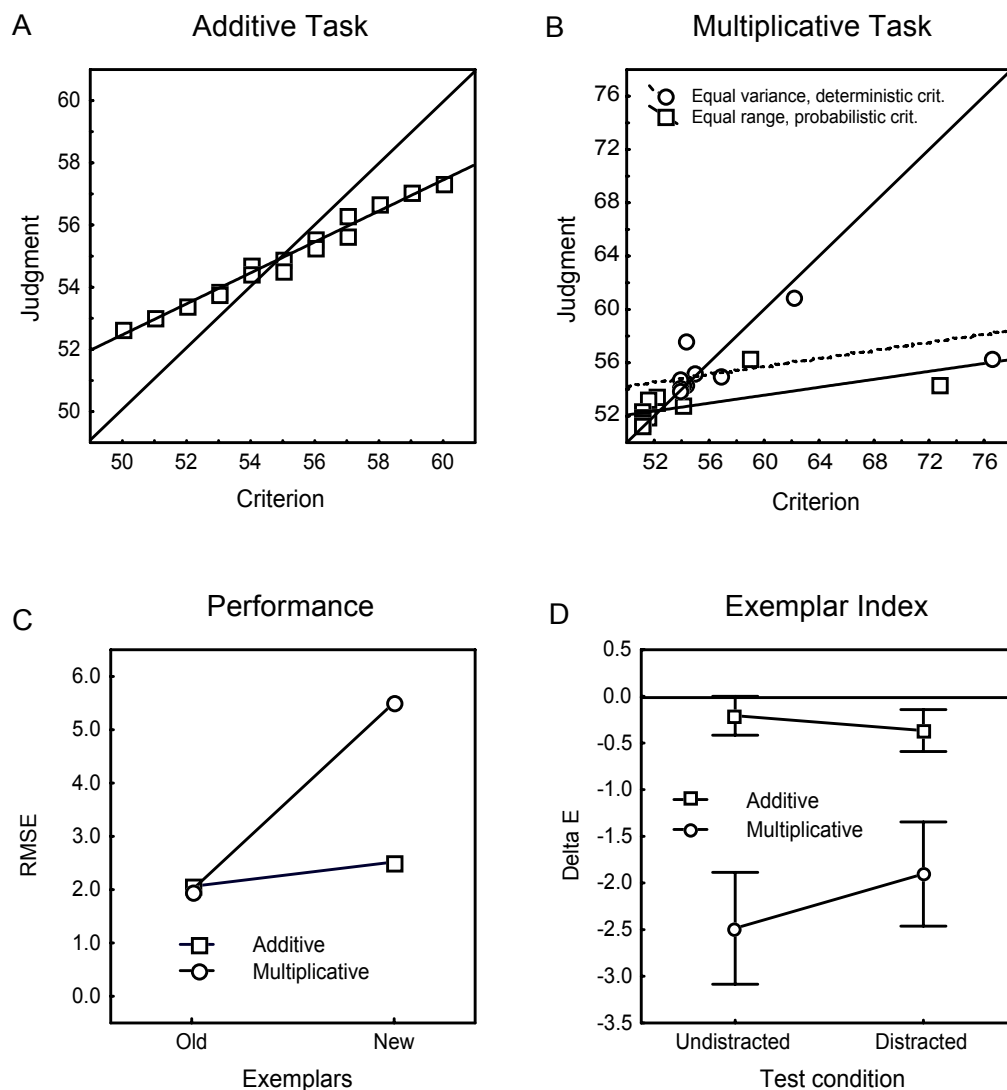


Figure 4: Panel A: Mean judgments in the additive task. Panel B: Mean judgments in the multiplicative task. Panel C: Performance measured as RMSE for old and new exemplars respectively compared between the additive and the multiplicative task. Panel D: Exemplar index in the additive and the multiplicative task in the undistracted and distracted test phase. Adapter from “Information Integration in Multiple Cue Judgment: A Division of Labor Hypothesis” by P. Juslin, L. Karlsson & H. Olsson, 2008. *Cognition*. Copyright by Elsevier.

In the additive task the exemplar index included 0, suggestive of cue abstraction. In the multiplicative task exemplar index was well below 0, indicating that exemplar memory had been the dominating process (Figure 4D).³ Consistent with the results suggested by the exemplar index, in the additive task, the additive cue abstraction model fitted the data better than the exemplar model

³There was a significant interaction between environment and undistracted or distracted test phase, in terms of exemplar index. However, since a priori the exemplar model only predicts a negative exemplar index, not its magnitude, this effect is not to be interpreted as a shift towards cue abstraction in the distracted test phase in the multiplicative task.

($r^2 = .95$ with $RMSD = .32$ versus $r^2 = .89$ with $RMSD = 0.45$), while the reverse was true in the multiplicative task ($r^2 = .46$ with $RMSD = 2.94$ versus $r^2 = .92$ with $RMSD = 0.49$).

7.4 Controlled vs. confounded training sequence (Juslin et al., 2008)

Experiment 2 in Juslin et al. (2008) also tested the predicted effect by manipulating the presentation order of the learning exemplars. Specifically, if the learning sequence is manipulated so that each successive trial only shows a

bug where only one cue has changed from the trial before, this should facilitate learning with cue abstraction. If you, for example, first observe a bug with cues [1, 1, 1, 0] and criterion 59 and then immediately thereafter observe a bug with cues [1, 1, 0, 0] and criterion 57, this invites the inference that the third cue, which is the only difference between the two exemplars, accounts for the difference of two units on the criterion, allowing you to abstract the weight of Cue 3. We administered two different learning sequences, one where there was always only one cue that changed from trial to trial and one sequence in which several cues always changed from trial to trial.

The hypothesis was that a controlled sequence, where only one cue changes between the successive trials, should improve learning in the additive task where people primarily rely on cue abstraction. By contrast, in the multiplicative task a controlled sequence should provide no benefit if people rely on exemplar memory. If anything, it should lead to poorer performance by inviting people to attempt futile attempts at cue abstraction, which, as implied by the constraints in Sigma, is virtually impossible in a multiplicative task.

The training phase was divided into blocks of 11 judgments. Judgment performance (RMSE between judgment and criterion) for the first two training blocks was taken to index the speed of learning and was entered into a factorial ANOVA that yielded a statistically significant interaction. Learning in the additive task was thus *facilitated* by the controlled sequence, while learning in the multiplicative task was *impaired* by the controlled sequence. This dissociation strongly suggests different cognitive processes in the two conditions. The 95% confidence intervals for the mean exemplar index ΔE in each cell included 0 only in the additive task with controlled sequence. As in Experiment 1, the cue abstraction model fitted data best in the additive task, while the reverse was true in the multiplicative task. The results from both Experiments 1 and 2 in Juslin et al. (2008) therefore clearly suggest that participants adapted to the multiplicative task by a shift to exemplar memory.

8 Stubborn resistance to shifts

8.1 Linear vs. non-linear cue-combination rule (Olsson, Enkvist et al., 2006)

Olsson, Enkvist et al. (2006) tested if a task with non-linear cue-criterion relations would induce a shift from cue abstraction to exemplar memory. The criterion in the nonlinear task involved a non-linear transformation of the linear criterion c_L in Eq. 1:

$$c_{NL} = ((-2 \cdot (c_L)^2)/5) + 44 \cdot c_L - 1150 \quad (7)$$

In Experiment 1 the idea was to compare judgments in an additive task with judgments in 1) a probabilistic non-linear task, 2) a probabilistic non-linear task with frequency manipulation and 3) a deterministic non-linear task with frequency manipulation. The frequency manipulation increased the presentation frequency of the extreme training exemplars that most clearly reveal the non-linear function in Eq. 7.

The hypothesis was that cue abstraction would be the dominating process only in the linear task. Inspecting Figure 5, plotting the mean judgments as a function of the criterion, reveals that performance is good in the linear condition but extremely poor in all three non-linear conditions. Although the task should be facilitated by making it deterministic and allowing the extreme exemplars to be shown especially often, only 60 % of the participants with the deterministic non-linear task with frequency manipulation had a significant correlation between their judgments and the criterion for the old exemplars.

The Exemplar Index was significantly separated from 0 in all three of the non-linear conditions, but includes 0 in the linear condition, suggesting that cue abstraction was the dominating process in the linear task. In the linear condition, the cue abstraction model provides the best quantitative fit. In all non-linear conditions the fit of both models was poor (r^2 between .13 and .26), basically because the judgment performance was very poor.

Given the complexity of the nonlinear task it may not appear surprising that the participants failed to abstract the underlying cue criterion relations. But the learning phase in the condition with a deterministic criterion involved no less than 20 presentations of the same 11 exemplars (i.e., the same four binary cues with the same criterion). This would seem to make exemplar memory both a viable and useful process to make accurate judgments, but evidently the participants were unable to shift to reliance on exemplar memory.

8.2 Prolonged training and explicit instructions (Olsson, Enkvist et al., 2006)

Is the non-linear task too difficult to accomplish during the 220 trial session? Experiment 2 in Olsson, Enkvist et al. investigated ways to improve learning and performance in the nonlinear task. The learning phase was therefore extended to 440 trials, altogether no less than 40 presentations of each unique exemplar! Further, the experiment was divided into two conditions; in one condition the participants received the same instructions as in Experiment 1, simply asking them to learn to predict the criterion, as in most studies of multiple cue judgment. In the other condition the instructions explicitly told the participants that there was no way in which the cue-criterion

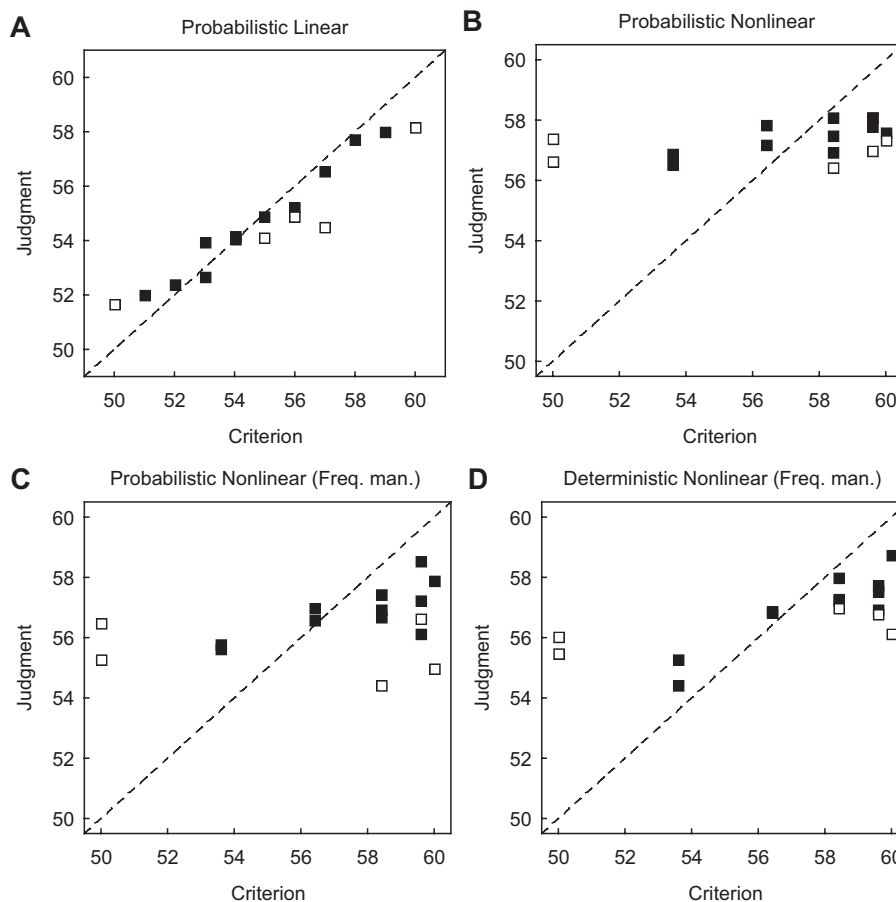


Figure 5: Mean judgment from the test phase of Experiment 1 plotted against the criterion. Filled squares are old exemplars, and open squares are new exemplars only presented in the test phase. Panel A: The probabilistic linear condition. Panel B: Probabilistic nonlinear condition. Panel C: Probabilistic nonlinear condition with frequency manipulation. Panel D: Deterministic nonlinear condition with frequency manipulation. Adapted from “Go with the flow: how to master a non-linear multiple-cue judgment task” by A-C. Olsson, T., Enkvist & P. Juslin, 2006. *Journal of Experimental Psychology: Learning, Memory and Cognition*. Copyright by the American Psychological Association.

relations can be abstracted and the only way to master the task is by memorizing the concrete exemplars. We call the conditions the *neutral condition* and the *exemplar condition*, respectively. One outcome could be that already with the neutral instruction 440 trials of learning would be sufficient for the shift to exemplar memory to materialize. The other possibility could be that the participants are unable to spontaneously perform this process shift also after 440 trials, and that we will only observe a shift to exemplar memory after providing explicit exemplar instructions.

The performance in terms of *RMSE* was significantly better in the exemplar instruction condition than in the neutral instruction condition, but performance in the neutral condition of Experiment 2 was not significantly improved from the corresponding deterministic condition in Experiment 1 with 220 trials, suggesting that despite 440

learning trials the participants were still unable to shift to exemplar memory (Figure 6). There was no difference in terms of Exemplar Index between the neutral and the exemplar conditions. The exemplar model provides the best fit in the exemplar condition (see Figure 7). This result not only validates the modeling, but also suggests that the participants were stubbornly resisting a shift to exemplar memory, unless they were explicitly told by instructions to perform a strategy shift.

9 Summary of the results

When the judgment involves a binary criterion exemplar memory plays a more important role than it does in judgments of a deterministic continuous criterion (Juslin, Olsson et al., 2003). Apparently, when providing feedback that is rich enough to allow abstraction of the underlying

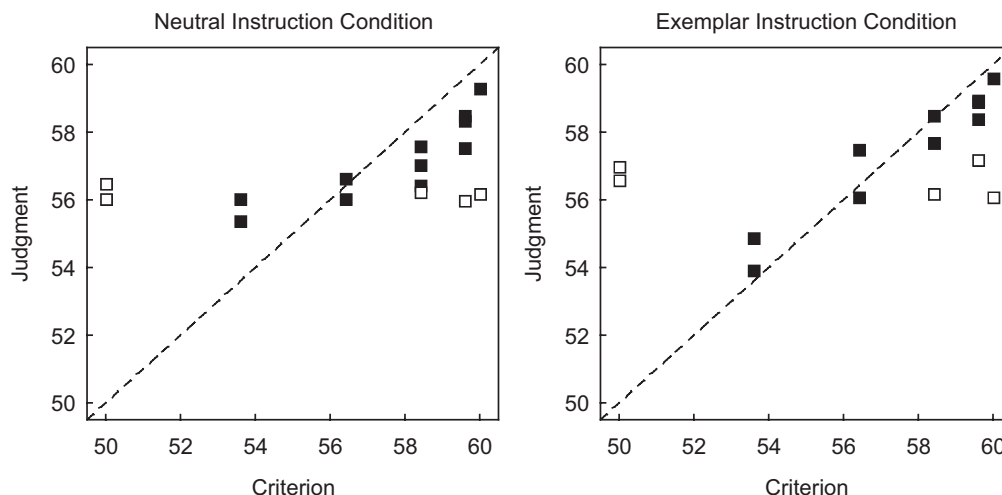


Figure 6: Mean judgment from test phase of Experiment 2 plotted against the criterion. Filled squares are old exemplars presented both under training and at test, and open squares are new exemplars only presented in the test phase. Left panel: The neutral instruction condition. Right panel: The exemplar instruction condition. Adapted from “Go with the flow: how to master a non-linear multiple-cue judgment task” by A-C. Olsson, T., Enkvist & P. Juslin, 2006. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Copyright by the American Psychological Association.

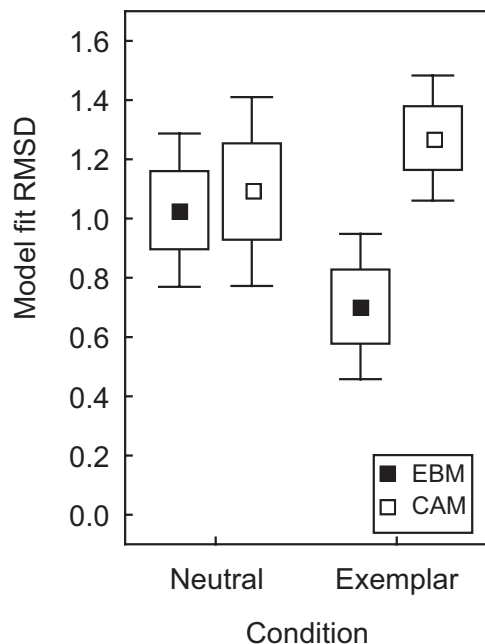


Figure 7: Model fit in terms of RMSD for each condition and model. (EBM = exemplar-based model, CAM = cue-abstraction model) Adapted from “Go with the flow: how to master a non-linear multiple-cue judgment task” by A-C. Olsson, T., Enkvist & P. Juslin, 2006. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Copyright by the American Psychological Association.

cue-criterion relations the participants favor cue abstraction. Moreover, when changing the cue-combination rule from an additive to a multiplicative function the participants are no longer able to add the impact of separate cues into accurate judgments, and therefore they seem to shift rapidly to exemplar memory (Juslin et al., 2008; see also Karlsson, Nyberg et al., 2007 for implications at the neural level).

The results in Olsson, Enkvist et al. (2006), however, provide clear evidence for a situation where such an adaptive shift apparently fails to materialize. Even though in the deterministic non-linear conditions there were a very limited number of unique exemplars (11) and the learning was extended to 440 trials the data reveal extremely poor learning in the non-linear task. In the contrast to the assumption of an automatic shift, only when the participants were told what strategy could be a viable alternative did they perform reasonably well.

10 General discussion

At a first glance, the poor learning observed in the non-linear tasks might seem rather surprising. Abilities to abstract statistical regularities in our environments or to learn mere stimulus-response associations have apparently not been important in the non-linear tasks reported above.

Our interpretation of the data reviewed in this article can be framed like this: 1) whenever possible, participants will try to abstract the underlying task structure as

captured by *cue abstraction* (a rule-bias, see Ashby et al., 1998; Juslin, Olsson et al., 2003). There are obvious advantages of having knowledge that is general enough to allow generalization also to other situations and which provides a readily communicated summary of the underlying task structure. 2) However, when attempts at cue abstraction are futile and participants do not manage to create stable representations of the linear slopes between cues and criterion, they will try the strategy of using *exemplar memory* instead. Exemplar memory allows a flexible adaptation to judgment tasks, since storing instances in memory should be less dependent on the task structure than cue abstraction. 3) However, if the initial attempt at using a strategy of exemplar memory leads to poor performance, the judge reverts back to the default strategy of cue abstraction and become stuck in impasse with extremely poor performance.

This interpretation implies that what happens early in the application of a specific strategy is crucial for determining the development of a stable strategy. (Additive) cue abstraction is obviously not appropriate for good performance in the non-linear task, as hypothesized by Sigma (Olsson, Enkvist et al., 2006; Juslin et al., 2008). But, why is the exemplar strategy rejected early in its application to this nonlinear judgment task?

At asymptotic training, exemplar memory allows perfect performance for the old exemplars in the deterministic nonlinear task. But when actually scrutinizing the nonlinear task (see Table 1) it becomes evident that early during learning, exemplar memory may not be encouraged by the non-linear task structure. After the first trials, having only a few exemplars stored in memory, perhaps with very dissimilar cue profiles but similar criterion values, might make integration of similar exemplars difficult while still aiming at good performance. For example, consider that Exemplar #2 and Exemplar #15 in Table 1 have cue profiles [1, 1, 1, 0] and [0, 0, 0, 1]. These two exemplars have diametrically opposed cue profiles (all four cues differ) but the same criterion (53.6). On the other hand, Exemplar #2 and Exemplar #6 in Table 1 have cue profiles [1, 1, 1, 0] and [1, 0, 1, 0] that are highly similar (only one cue differs), but rather large difference in the criterion (53.6 & 59.6, respectively). This complex similarity structure may dissuade early attempts at applying an exemplar strategy.

We conclude that the results reviewed in this article speak strongly against the assumption of parallel and obligatory encoding of exemplars, as envisioned by Logan (1988), at least in the context of multiple-cue judgments. We also believe that the results speak less in favor of a mechanism that is a parallel and automatic side effect of experience (e.g., Bourne et al., 1999; Erickson & Kruschke, 1998; Logan, 1988; Palmeri, 1997; Rickard, 1997, 2004), but more in favor of a mechanism where

memory traces for later recollection and usage as a basis for judgment demands a strategic choice on behalf of the participant. This alternative explanation of the nature of the interplay between exemplar memory and other processes may thus be best understood as *strategy shifts*, contingent on the structure of the task.

There are other lines of research that appear to support our way of reasoning. To date there are a number of empirical studies demonstrating how explicit processes can hinder effective learning with implicit processes (e.g. Fletcher et al., 2005; Reber, 1976). With the first words of the title intriguingly chosen as “*On the benefits of not trying...*” Fletcher et al. state that explicit processes might actually *suppress* implicit learning. They report findings where explicit intention to learn a difficult sequence impaired implicit learning of the same sequence.

Moreover, Sloutsky and Fischer (2004) suggests that children have better memory performance than adults after learning inductive categorization tasks, since they have not yet reached the developmental criteria for inducing category-specific information, and are hence constrained to use a memory-based strategy. This invites the counterintuitive hypothesis that children might perform better than adults in the non-linear task reported in this article.

10.1 Conclusions

Conceiving exemplar memory as a strategy to be chosen rather than as an automatic side effect of learning is an important step towards understanding the cognitive basis of multiple-cue judgment. Further research endeavors are necessary to get firmer support for this view. There are not that many published multiple-cue judgment (or categorization) studies with *healthy participants* that report a task where virtually no learning was observed (but see Smith, Redford, Gent & Washburn, 2005). Nonetheless, the results reviewed in this article suggests that learning in multiple-cue judgment, including shifts of the representation that is input to the process, may have more to do with active and controlled problem solving, than with automatic side effects of experience as captured by associative mechanisms.

References

- Aha, D. W. (1997). Lazy learning. *Artificial Intelligence Review*, 11, 7–10.
- Anderson, J. R., & Fincham, J. (1996). Categorization and sensitivity to correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 259–277.

- Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442–481.
- 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.
- Brehmer, B. (1994). The psychology of linear judgment models. *Acta Psychologica*, *87*, 137–154.
- Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgments of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman, (Eds.) *Heuristics and biases: The psychology of intuitive judgement*, pp. 120–138. New York: Cambridge University Press.
- Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. San Diego: Academic Press, Inc.
- DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 968–986.
- Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*, 951–978.
- Einhorn, J. H., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Regression models and process tracing analysis. *Psychological Review*, *86*, 465–485.
- Ellis, R., & Humphreys, G. W. (1999). *Connectionist Psychology: A Text with Readings*. London: Psychology Press.
- Enkvist, T., Newell, R. B., Juslin, P., & Olsson, H. (2006). On the role of casual intervention in multiple-cue judgment: Positive and negative effects on learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 163–179.
- Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107–140.
- Estes, W. K. (1994). *Classification and cognition*. New York: Oxford University Press.
- Fischbein, E., Deri, M., Nello, M., & Marino, M. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*, 3–17.
- Fletcher, P. C., Zafiris, O., Frith, C. D., Honey, R. A. E., Corlett, P. R., Zilles, K., & Fink, G. R. (2005). On the benefits of not trying: brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex*, *15*, 1002–1015.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Haider, H., Frensch, P. A., & Joram, D. (2005). Are strategy shifts caused by data-driven processes or by voluntary processes? *Consciousness and Cognition*, *14*, 495–519.
- Hammond, K. R., & Stewart, T. R. (Eds.) (2001). *The essential Brunswik: Beginnings, explications, applications*. Oxford: Oxford University Press.
- Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411–428.
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1–55.
- Juslin, P., Jones, S., Olsson, H., & Winman, A. (2003). Cue abstraction and exemplar memory in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 924–941.
- Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple-cue judgment: A division of labor hypothesis. *Cognition*, *106*, 259–298.
- Juslin, P., Olsson, H., & Olsson, A-C. (2003). Exemplar effects in categorization and multiple-cue judgment. *Journal of Experimental Psychology: General*, *132*, 133–156.
- Juslin, P., & Persson, M. (2002). PROBabilities from Exemplars (PROBEX): A "lazy" algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, *26*, 563–607.
- Karlsson, L., Juslin, P., & Olsson, H. (2007). Adaptive changes between cue abstraction and exemplar memory in a multiple-cue judgment task with continuous cues. *Psychonomic Bulletin and Review*, *14*, 1140–1146.
- Karlsson, L., Nyberg, L., Juslin, P., & Olsson, H. (2007). Distinct neural correlates underlie multiple-cue judgment depending on the cue-combination rule. Manuscript submitted for publication.
- Logan, D. G. (1988). Towards an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.
- Meeter, M., Myers, C. E., Shohamy, D., Hopkins, R.O., & Gluck, M. A. (2006). Strategies in probabilistic categorization: Results from a new way of analyzing performance. *Learning and Memory*, *13*, 230–239.
- Mellers, B. A. (1980). Configurality in multiple-cue probability learning. *American Journal of Psychology*, *93*, 429–443.

- Nilsson, H., Olsson, H., & Juslin, P. (2005). The cognitive substrate of subjective probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 600–620.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 104–114.
- Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.
- Nosofsky, R. M., & Johansen, M. K. (2000). Exemplar-based accounts of “multiple-system” phenomena in perceptual categorization. *Psychonomic Bulletin and Review*, *7*, 375–402.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.
- Olsson, A.-C., Enkvist, T., & Juslin, P. (2006). Go with the flow! How to master a non-linear multiple cue judgment task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1371–1384.
- Olsson, A.-C., Juslin, P., & Olsson, H. (2006). Multiple cue judgment in individual and dyadic learning. *Journal of Experimental Social Psychology*, *42*, 40–56.
- Olsson, H., & Poom, L. (2005). The noisy cue abstraction model is equivalent to the multiplicative prototype model. *Perceptual and Motor Skills*, *100*, 819–820.
- Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 324–354.
- Pothos, E. M., & Bailey, T. M. (2000). The role of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 847–862.
- Reber A. S. (1976). Implicit learning of synthetic languages: the role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 88–94.
- Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*, 1–41.
- Rickard, T. C. (1997). Bending the power law: a CMPL theory of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, *126*, 288–311.
- Rickard, T. C. (2004). Strategy execution in cognitive skill learning: an item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 65–82.
- Roussel, J.-L., Fayol, M., & Barrouillet, P. (2002). Procedural vs. direct retrieval strategies in arithmetic: A comparison between additive and multiplicative problem solving. *European Journal of Cognitive Psychology*, *14*, 61–104.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.
- Sieck, W. R., & Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1003–1021.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1–19.
- Sloutsky, V. M., & Fisher, A. V. (2004). When development and learning decrease memory — evidence against category-based induction in children. *Psychological Science*, *15*, 553–558.
- Smith, J. D., Redford, J.S., Gent, L. C. & Washburn, D.A. (2005). Visual search and the collapse of categorization. *Journal of Experimental Psychology: General*, *134*, 443–460.
- Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review*, *99*, 3–21.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organisation of memory*. London: Academic Press.