Blending and Choosing Within One Mind: Should Judgments Be Based on Exemplars, Rules, or Both?

Stefan M. Herzog (herzog@mpib-berlin.mpg.de)
Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94
14195 Berlin, Germany

Bettina von Helversen (bettina.vonhelversen@unibas.ch)
Department of Psychology, University of Basel, Missionsstrasse 62a
4055 Basel, Switzerland

Abstract

Accurate judgments and decisions are crucial for success in many areas of human life. The accuracy of a judgment or decision depends largely on the cognitive process applied. In research on judgment, decision making, and categorization, two kinds of cognitive processes have often been contrasted: exemplar-based processes, which use similarity to previously encountered items to make judgments, decisions, and categorizations, and rule-based processes, which use abstracted cue knowledge. Although most cognitive models of judgment and decision processes assume that people rely on both processes, they differ in whether they assume that one process is selected or that both processes are blended into a single response. The present research takes a functional perspective and investigates what kind of interaction between the two processes leads to accurate responses. Based on cross-validated simulations in real-world domains, it shows that blending rule- and exemplar-based processes generally leads to better judgments than does choosing between them, suggesting that the default strategy should be a blend of both processes, which is abandoned only when feedback justifies it.

Keywords: accuracy; multiple-cue judgments; decision making; categorization; exemplar models; rules; cognitive models; mixtures of experts; simulation.

Introduction

Judging quantities, making decisions, and categorizing items are crucial elements of successful human behavior. A vast and diverse literature in cognitive science and judgment and decision making has investigated how people achieve these tasks (e.g., Ashby & Maddox, 2005; Gigerenzer, Hertwig, & Pachur, 2011; Kruschke, 2008; Payne, Bettman, & Johnson, 1993). The many different models and strategies proposed can be broadly classified into two categories with reference to the cognitive processes they assume: exemplar-based processes, which use similarity to previously encountered items to make judgments, decisions, and categorizations, and rule-based processes, which use abstracted cue knowledge (Hahn & Chater, 1998).

Extensive research has compared the proposed models' ability to describe human behavior. Furthermore, the performance of judgment and decision making strategies in predicting real-world criteria has been thoroughly investigated (e.g., Gigerenzer et al., 2011; Todd, Gigerenzer, & the ABC Research Group, 2012).

To our knowledge, however, research in cognitive science and judgment and decision making has not previously investigated what kind of interaction between exemplar- and rule-based processes leads to accurate judgments, decisions, and categorizations: relying on just one of the two processes or using both? If both are considered, is it better to choose between them depending on the structure of the task, for instance (Rieskamp & Otto, 2006), or to blend them into a joint response? This paper presents first answers to these questions.

A functional perspective on the interaction between exemplar- and rule based processes may be useful for at least three reasons. First, examining cognitive models' ability to predict external real-world criteria goes a step further than comparing their ability to describe human behavior in idealized laboratory tasks, by adding a further evaluation criterion. If one class of cognitive models were superior to another in terms of predictive performance, this would make them more attractive as plausible models of human behavior (Chater & Oaksford, 1999). Second, many cognitive models are inspired by or share similarities with models from research fields interested in predictive performance (such as statistics, artificial intelligence, computer science, and machine learning; see e.g., Jäkel, Schölkopf, & Wichmann, 2009; Marling, Sqalli, Rissland, Munoz-Avila, & Aha, 2002), and a functional perspective provides a common ground that serves to re-connect cognitive models with such fields. Third, knowledge of how to profit from the complementary strengths of the two processes could offer prescriptions for improving human judgment, decision making, and categorization by instructing decision makers on when and how to use the two processes.

Models of Judgment, Decision Making, and Categorization

There are two general approaches to modeling human cognition. First, single general-purpose models have been proposed (e.g., Lee & Cummins, 2004). For instance, judgment and categorization models assume either only exemplar-based (e.g., Juslin & Persson, 2002; Kruschke, 1992) or only rule-based processes (e.g., Ashby & Gott, 1988; Brehmer, 1994). Second, toolbox approaches have been proposed. These assume that people draw on multiple, different processes to solve the same task (e.g., Gigerenzer & Selten, 2001). The toolbox approach posits that people adaptively select a tool (i.e., strategy) likely to succeed in the task at hand from a repertoire of strategies: the "toolbox" (Gigerenzer & Selten, 2001; Payne et al., 1993; Rieskamp & Otto, 2006; Scheibehenne, Rieskamp, & Wagenmakers, 2013). Toolbox approaches have gained popularity particularly in decision making (e.g., Gigerenzer & Selten, 2001; Rieskamp & Otto, 2006). Yet also in categorization and judgment research, it is frequently assumed that people chose the process that is better suited to solving a task (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Juslin, Karlsson, & Olsson, 2008; Nosofsky, Palmeri, & McKinley, 1994; von Helversen & Rieskamp, 2008). For example, COVIS assumes that similarity-based and rule-based processes "race" for an answer, with the faster one determining the response (Ashby et al., 1998).

Although toolbox approaches often assume competition between processes, it is also possible that the processes cooperate. Hybrid or blending models assume that, instead of "choosing" a process for a task, two or more processes are executed simultaneously and their responses are integrated. For instance, the categorization model ATRIUM (Erickson & Kruschke, 1998) combines both exemplar- and rule-based processes. Inspired by the "mixtures-of-experts" approach from machine learning (Jacobs, Jordan, Nowlan, & Hinton, 1991), ATRIUM assumes that people have two "experts" in their mind: an exemplar-based and a rule-based one, whose outputs are processed by a gating mechanism. This gating mechanism can "choose" between these modules or "blend" their outputs by averaging their responses. In addition, ATRIUM can learn to rely more strongly on the more successful module (in terms of the probability of choosing or weighted averaging)—either for the whole task or depending on the item presented (i.e., depending on its location in psychological space). Modeling and experimental investigations support ATRIUM's assumption that exemplar- and rule-based processes simultaneously influence how humans categorize (e.g., Erickson & Kruschke, 1998; Hahn, Prat-Sala, Pothos, & Brumby, 2010). There is also evidence for such simultaneous influence in the domain of multiple-cue judgments (von Helversen, Herzog, & Rieskamp, in press).

Blending and Choosing Within One Mind

The combination of judgments or decisions from different sources is a vibrant topic in research fields such as psychology, judgment and decision making, cognitive science, statistics, artificial intelligence (AI), machine learning, biology, and economics (e.g., Krause, Ruxton, & Krause, 2010; Kuncheva, 2004; Larrick, Mannes, & Soll, 2012; Lee, Zhang, & Shi, 2011; Marling et al., 2002). Combining diverse sources (e.g., forecasts from different experts) generally improves accuracy because different sources often compensate for each other's shortcomings. Depending on the circumstances, either choosing between ("competition") or blending different sources ("cooperation") may lead to better performance.

On the one hand, choosing a specific strategy allows the overall decision process to be adapted to environmental regularities and thus facilitates good performance (e.g., Todd et al., 2012). On the other hand, "blending" (i.e., averaging) different sources can often improve accuracy because errors of different signs cancel each other out. This "wisdom of crowds" phenomenon (Surowiecki, 2004) has recently also been applied to individual minds (e.g., Herzog & Hertwig, 2009, 2013; Vul & Pashler, 2008). Combining exemplar- and rule-based processes can be seen as an implicit "crowd within," where the two processes constitute two "experts" in one mind that either compete or cooperate in giving a response. To the extent that exemplar- and rule-based processes complement each other in the errors they commit, combining them may be a successful strategy (Herzog & von Helversen, 2013).

In the following simulation study, we compare the merits of single purpose models, a competitive toolbox approach, and a cooperative toolbox approach. We focus on exemplar-based and rule-based processes as examples of distinctive cognitive processes because of the prominent distinction between the two in the cognitive literature (Ashby et al., 1998; Hahn & Chater, 1998; Nosofsky et al., 1994; Persson & Rieskamp, 2009).

Different Levels of Interaction: Task or Item

Besides differentiating between choosing (competition) and blending (cooperation) of cognitive processes, we also consider on which level the interaction takes place: the task or item level. In the ecological rationality and adaptive toolbox approach (Todd et al., 2012), it is (implicitly) assumed that the selection of strategies happens on the task level—that is, that all the decisions within the same task are solved using the same strategy (once learning has completed). However, strategy selection (or integration) can also happen on the item level—that is, some items may be better solved by a rule, whereas others require memorization (Nosofsky et al., 1994). To account for this level of interaction, we compared competition and cooperation on the task and the item level.

Simulation Study: Should Judgments Be Based on Exemplars, Rules or Both?

We investigated the performance of different ways to use exemplar- and rule-based processes in predicting a continuous criterion based on multiple cues. To this end, we conducted cross-validated simulations, informed by ATRIUM's (Erickson & Kruschke, 1998) cognitive architecture, in five real-world domains. We addressed the following three questions. First, is it better to be equipped with both exemplar- and rule-based processes or is one process enough to achieve accurate judgments? Second, if both processes are used, is it better to choose between them

Table 1: Characteristics of the real-world datasets (adapted from Table 1 in Dana & Dawes, 2004). N = number of cases, k = number of cues, ρ = correlation between target variable and predicted values from a multiple linear regression, **v** Vector = zero-order correlation between target variable and cues, $\varnothing r_{xixj}$ = mean correlation among cues.

Dataset	N	k	ρ	**v** Vector	$\varnothing r_{xixj}$
Abalone	4,177	7	.73	.63 .58 .56 .56 .54 .50 .42	.89
NFL	3,057	10	.54	.46 .43 .37 .34 .33 .27 .21 .07 .05 .05	.21
ABC	955	5	.35	.32 .20 .06 .04 .02	.08
NES	1,910	6	.35	.26 .17 .15 .15 .13 .12	.11
WLS	6,385	5	.20	.13 .11 .10 .10 .10	.15

(competition) or to blend them (cooperation)? Third, for either choosing between or blending the two processes, is it better to treat all items the same (i.e., integration on the task level) or to treat individual items differently (i.e., integration on the item level)? Item-level integration implies choosing between the processes for each item (in the competitive approach) or weighting the two processes differently for each item when blending (in the cooperative approach).

Datasets

We analyzed datasets previously used to compare the performance of proper and improper linear models (Dana & Dawes, 2004). The datasets pertain to five domains: biology, sports, public opinion, political sentiment, and occupational prestige. In all datasets, a continuous target variable was predicted by several cues. For instance, the ABC dataset was derived from a 2002 poll of 955 U.S. households. Respondents' confidence that Osama bin Laden would be captured or killed was predicted by five cues, including the respondent's age, education, gender, and patriotism. See Table 1 for details of the statistical structure.

Cognitive Models

Exemplar Model To represent an exemplar-based judgment process, we used an exemplar model for multiple-cue judgments (Juslin et al., 2008). The model assumes that judgments are based on the similarity to exemplars stored in memory, where the judgment is an average of the criterion values of the stored exemplars weighted by their similarity to the target item. We used a simplified exemplar model with one single free parameter determining the similarity gradient (see von Helversen & Rieskamp, 2008).

Rule Model To represent a rule-based process, we used a multiple linear regression model. Such models have been widely used to model human judgment (Brehmer, 1994); they assume that judgments can be understood as the sum of weighted cue values. The model has a free parameter for every cue plus an intercept.

Simulation Setup

For each simulation run, we randomly drew a learning sample and a test sample. We then fitted the free parameters of the exemplar and the rule model to the learning sample—minimizing the root mean square error (RMSE) between model predictions and criterion values—and used the estimated parameter values to make predictions for the items in the test sample (for six different strategies described below). We measured estimation accuracy in the test sample using the RMSE between the model's predictions and the criterion values, a commonly used measure of absolute goodness of fit. Seven different sizes of learning samples were used (20, 40, 60, 80, 100, 200, and 500 items) to vary the amount of experience with a domain; all test samples consisted of 250 items. For each dataset and each of the sizes of learning samples, we ran the simulation 1,000 times and averaged the results.

Using the Rule and Exemplar Models

We tested six strategies for using rule- and exemplar-based processes to make predictions for the test sample.

"Exemplar Model" and "Rule Model" The first two strategies used just one of the two processes exclusively.

"Choosing-Task" and "Choosing-Item" The third and fourth strategy chose either the exemplar or the rule model.

On the task level, "choosing-task" selected in each simulation run the model that was superior in the learning sample and used it for all items in the test sample. To account for differences in model complexity, we used the Bayesian Information Criterion as a selection criterion.

On the item level, "choosing-item" selected in each simulation run *and for each item in the test sample* the model that was more likely to be superior for this particular test item—based on the performance on similar items in the learning sample. Specifically, for each test item we calculated the RMSE that the exemplar and the rule model had on similar items in the learning sample (i.e., we weighted the RMSE values of each training item using the similarity gradient of the exemplar model). The process with the lower weighted RMSE was then selected and its prediction for this test item was used.

"Blending-Average" and "Blending-Item" The fifth and sixth strategy blended the outputs of the exemplar and the rule model to make a joint prediction.

On the task level, "blending-average" computed for each test item the arithmetic mean of the predictions of the rule and the exemplar model.

On the item level, "blending-item" used in each simulation run *and for each item in the test sample* a

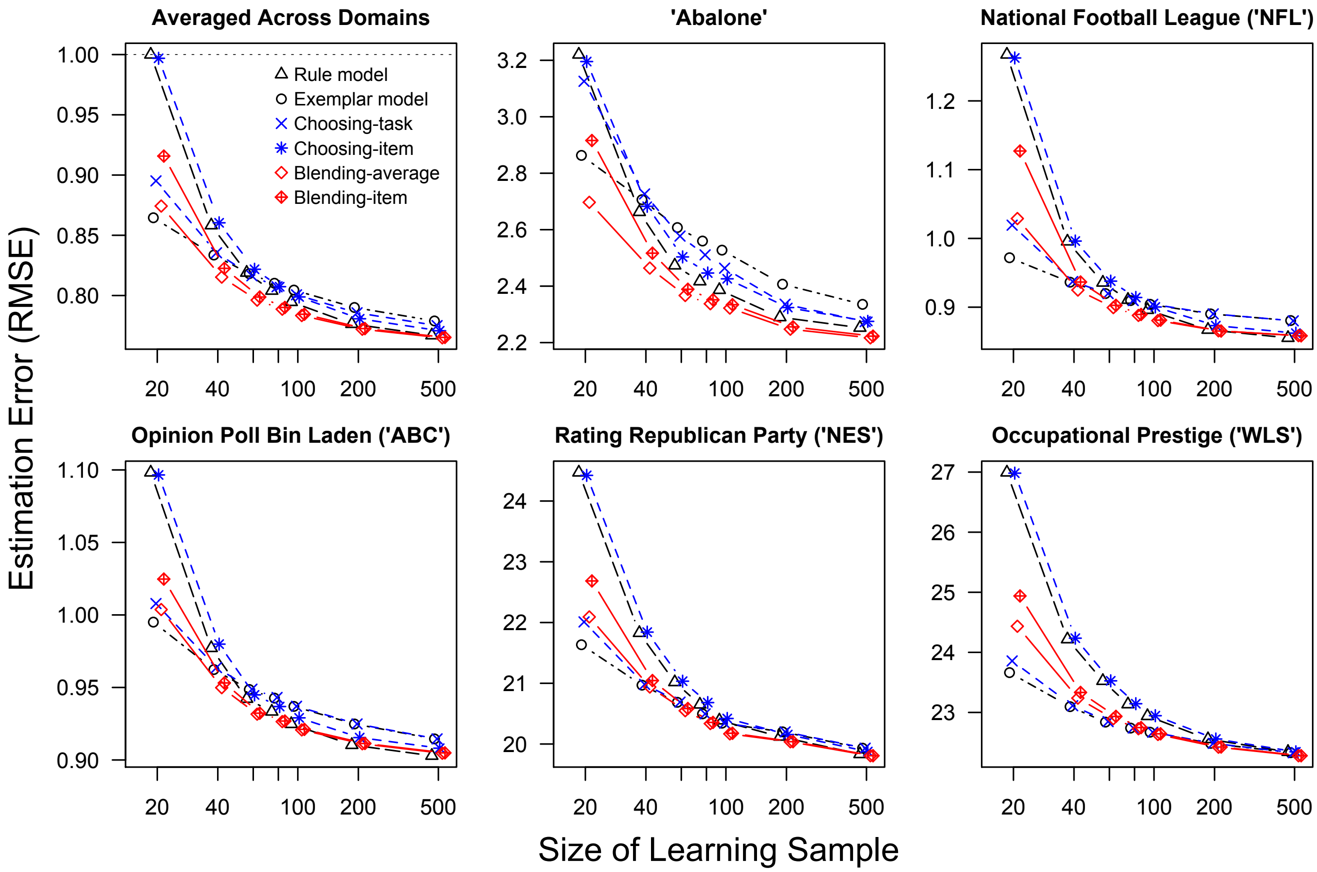

Figure 1: Cross-validated estimation accuracy (Root Mean Squared Error, RMSE) of six strategies in five domains (for learning samples of different sizes). The upper left panel averages the normalized data across domains; the RMSE values were divided by the largest average RMSE value in each domain. The strategies are explained in the text.

weighted average of both models' predictions—using the same similarity-weighted RMSEs as in "choosing-item." The item-specific weight for the exemplar model was calculated as the proportion of the rule model's weighted RMSE relative to the sum of both models' weighted RMSEs (i.e., the worse the rule model, the larger the weight on the exemplar model).

Results & Discussion

Figure 1 shows the generalization performance of the different strategies as a function of the size of the learning sample for the five domains. Because the datasets differed in their range of criterion values, which in turn affected the scale of the RMSE, it was necessary to normalize the RMSEs before aggregating them across datasets. To this end, we divided each RMSE by the largest average RMSE value within the respective domain, so that each RMSE value could be understood as the relative increase in fit. We then constructed a summary learning curve by averaging the normalized RMSEs across the five domains (see Figure 1, upper left panel).

Four results are noteworthy. First, "blending-average" was generally more accurate than either the exemplar or the rule model; the exemplar model was somewhat better than the averaged predictions of both models only for very small learning samples (i.e., 20 items). Second, "blending-average" was generally more accurate than choosing the better model based on its performance in the respective learning sample ("choosing-task"), although choosing was slightly better for very small learning samples (i.e., 20 items) in two of the five datasets. Third, when choosing or blending, it did not pay off to tune one's use of the models to the type of item. Weighting both processes when blending ("blending-item") was less or equally accurate than was giving them equal weights ("blending-average"); similarly, choosing the process depending on the item ("choosing-item") was less or equally accurate than was using the same process for all items ("choosing-task"). Fourth, the differences between strategies decreased as the size of the learning samples increased.

Let us now answer the three questions motivating this simulation. First, in the datasets we investigated, it was generally better to be equipped with both exemplar- and rule-based processes than with just one of the two processes. Second, if both processes were used, it was generally better to blend them than to choose between them. Third, when choosing between or blending the two processes, it was generally better to treat all items the same (and not to choose or blend, respectively, depending on the type of item; i.e., depending on how much "expertise" the exemplar- and rule-based processes had about a specific part of the psychological space).

General Discussion

Many cognitive models of judgment, decision-making, and categorization assume that people can use both exemplar- and rule-based processes (e.g., Erickson & Kruschke, 1998). Yet it remained unclear whether using both processes provides a performance advantage over using just one process and, when both processes are available, whether it is better to choose one process depending on the task (i.e., competitive toolbox approach) or to blend their responses (i.e., cooperative toolbox approach). Our simulations in the domain of multiple-cue judgments suggest that combining the two processes (either by choosing between or blending them) leads to better judgments than does relying on just one of them, and that a simple blend (i.e., equal weighting) of both processes leads to accurate judgments. This latter point is consistent with the success of naïve equal weighting strategies (e.g., Dawes, 1979). In another set of simulations, we investigated the combination (i.e., choosing or blending) of exemplar- and rule-based processes in the context of making categorizations (using 38 machine learning benchmark datasets; Herzog & von Helversen, 2013). Further broadening the scope of the present analysis, we found that blending the outputs of an exemplar- and a rule-based process led to successful categorizations.

Our results resonate with research in AI and machine learning that demonstrates how combining different representations is often beneficial (Kuncheva, 2004; Marling et al., 2002). More specifically, our results suggesting that combining exemplar- and rule-based processes can often increase accuracy in human cognition dovetail nicely with the successful combination of case-based and rule-based reasoning systems in AI (e.g., Marling et al., 2002; Prentzas & Hatzilygeroudis, 2007).

Besides the general question of whether exemplar- and rule-based processes should be "blended" or "chosen" among, our simulations suggest that it does not pay off to tune one's use of exemplar- and rule-based processes to the type of item one wants to generalize to. This conclusion seems inconsistent with empirical studies suggesting that participants successfully choose between processes in categorization tasks (e.g., Erickson, 2008). Yet these experimental tasks may be unrepresentative of real-world situations. In many experimental studies—especially in categorization research—there is little (or no) doubt about which process is better suited to solving the whole task (or responding to a specific item), and a participant can thus learn to choose between or differentially use the two processes. We speculate that deviating from a simple blending strategy is generally worthwhile only in domains in which one process is clearly superior to the other, both processes make similar errors, and this statistical structure can be ascertained with enough confidence (see Soll & Larrick, 2009). However, we would argue that this is typically not the case in real-world domains. It would thus seem prudent that human judges and decision makers, as modeled, for example, by ATRIUM (Erickson & Kruschke, 1998), start with a simple blend of both processes and deviate from this approach (e.g., by choosing or item-specific tuning) only when feedback justifies it.

Why is combining exemplar- and rule-based processes so successful in multiple-cue judgment tasks? The use and the performance of exemplar- and rule-based processes in multiple-cue judgment tasks seems to depend on the statistical structure of the task—in particular, the functional relation between cues and criteria (Juslin et al., 2008; von Helversen & Rieskamp, 2008). If the criterion can be approximated by a linear additive combination of the cues, rule-based processes predominate. In multiplicative tasks, by contrast, exemplar-based processes perform better and are used more frequently. Simulations using artificially created domains (Herzog & von Helversen, 2013) suggest that the five real-world domains we analyzed in the present simulations represent a mixture of these two kinds of statistical structures (i.e., additive and multiplicative). Consequently, neither of the two processes in isolation was able to capture their statistical structure. To the extent that this result generalizes to decision making and categorization, it suggests one reason why people are equipped with and use both exemplar- and rule-based processes: because only a combination of the two allows people to make successful judgments, decisions, and categorizations in the real world.

Acknowledgments

We thank the Swiss National Science Foundation for a grant to the first author (100014_129572/1).

References

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. (1998). A neuropsychological theory of multiple systems in category learning, *105*, 442–481.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 33–53.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178.

Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica*, *87*, 137–154.

Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, *3*, 57–65.

Dana, J., & Dawes, R. M. (2004). The superiority of simple alternatives to regression for social science predictions. *Journal of Educational and Behavioral Statistics*, *29*, 317–331.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.

Erickson, M. A. (2008). Executive attention and task switching in category learning: Evidence for stimulus-dependent representation. *Memory & Cognition*, *36*, 749–761.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107–140.

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford: Oxford University Press.

Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT Press.

Hahn, U., & Chater, N. (1998). Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition*, *65*, 197–230.

Hahn, U., Prat-Sala, M., Pothos, E. M., & Brumby, D. P. (2010). Exemplar similarity and rule application. *Cognition*, *114*, 1–18.

Herzog, S. M., & Hertwig, R. (2009). The wisdom of many in one mind: Improving individual judgments with dialectical bootstrapping. *Psychological Science*, *20*, 231–237.

Herzog, S. M., & Hertwig, R. (2013). The crowd-within and the benefits of dialectical bootstrapping: A reply to White and Antonakis (2013). *Psychological Science*, *24*, 117–119.

Herzog, S. M., & von Helversen, B. (2013). *The benefits of combining cognitive processes*. Manuscript in preparation.

Jacobs, R. A., Jordan, M., Nowlan, S., & Hinton, G. (1991). Adaptive mixtures of local experts. *Neural Computation*, *3*, 79–87.

Jäkel, F., Schölkopf, B., & Wichmann, F. A. (2009). Does cognitive science need kernels? *Trends in Cognitive Sciences*, *13*, 381–388.

Juslin, P., Karlsson, L., & Olsson, H. (2008). Information integration in multiple cue judgment: A division of labor hypothesis. *Cognition*, *106*, 259–298.

Juslin, P., & Persson, M. (2002). PROBabilities from Exemplars (PROBEX): A “lazy” algorithm for probabilistic inference from generic knowledge. *Cognitive Science*, *26*, 563–607.

Krause, J., Ruxton, G. D., & Krause, S. (2010). Swarm intelligence in animals and humans. *Trends in Ecology and Evolution*, *25*, 28–34.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

Kruschke, J. K. (2008). Models of categorization. In R. Sun (Ed.), *The Cambridge handbook of computational psychology*. New York, NY: Cambridge University Press.

Kuncheva, L. (2004). *Combining pattern classifiers: Methods and algorithms*. Hoboken, NJ: John Wiley & Sons.

Larrick, R. P., Mannes, A. E., & Soll, J. B. (2012). The social psychology of the wisdom of crowds. In J. I. Krueger (Ed.), *Frontiers in social psychology: Social judgment and decision making*. New York, NY: Psychology Press.

Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, *11*, 343–352.

Lee, M. D., Zhang, S., & Shi, J. (2011). The wisdom of the crowd playing The Price Is Right. *Memory & Cognition*, *39*, 914–923.

Marling, C., Sqalli, M., Rissland, E. L., Munoz-Avila, H., & Aha, D. (2002). Case-based reasoning integrations. *AI Magazine*, *23*, 69–86.

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

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.

Persson, M., & Rieskamp, J. (2009). Inferences from memory: Strategy- and exemplar-based judgment models compared. *Acta Psychologica*, *130*, 25–37.

Prentzas, J., & Hatzilygeroudis, I. (2007). Categorizing approaches combining rule-based and case-based reasoning. *Expert Systems*, *24*, 97–122.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.

Scheibehenne, B., Rieskamp, J., & Wagenmakers, E.-J. (2013). Testing adaptive toolbox models: A Bayesian hierarchical approach. *Psychological Review*, *120*, 39–64.

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 780–805.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. Garden City, NY: Doubleday.

Todd, P. M., Gigerenzer, G., & the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. Oxford, UK: Oxford University Press.

von Helversen, B., Herzog, S. M., & Rieskamp, J. (in press). Haunted by a Doppelgänger: Irrelevant facial similarity affects rule-based judgments. *Experimental Psychology*.

von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General*, *137*, 73–96.

Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *19*, 645–647.