Learning in Dynamic Probabilistic Environments: A Parallel-constraint Satisfaction Network-model Approach

Marc Jekel[1], Andreas Glöckner[1,2], & Arndt Bröder[3]

Draft, Friday 10th July, 2015

[1]University of Göttingen

[2]Max Planck Institute for Research on Collective Goods

[3]University of Mannheim

Word count.

Keywords. learning, decision making, parallel-constraint satisfaction network model, strategy-selection-learning theory, rationality

Note. This work was supported by Grant GL 632/3-1 and BR 2130/8-1 to Andreas Glöckner and Arndt Bröder from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program "New Frameworks of Rationality" (SPP 1516).

Correspondence should be addressed to. Marc Jekel, Georg-August-Universität Göttingen, Georg-Elias-Müller-Institut für Psychologie, Abt. 9: Psychologische Diagnostik, Urteilen und Entscheiden, Goßlerstraße 14, D-37073 Göttingen, Phone: +49-(0)551/3920624, E-mail: mjekel@psych.uni-goettingen.de.

Abstract

Learning is a crucial requirement for efficient decision making in the real world. Learners are able to adapt to decision-environments and also respond to changes in those environments. Previous research in the field of judgment and decision making has often ignored aspects of learning by focusing on one-shot decisions and static environments only. In this article, we develop an integrative model for decision making and learning by extending previous work on parallel constraint satisfaction networks with an algorithms of backward error-propagation learning. The Parallel Constraint Satisfaction Theory for Decision Making and Learning (PCS-DM-L) conceptualizes decision making as process of coherence structuring in which learning is achieved by adjusting network weights from one decision to the next. PCS-DM-L predicts that individuals adapt to the environment by gradual changes in cue weighting. This prediction competes with the fundamental assumption of adaptive decision making approaches, assuming that learning and adaptation to environmental changes takes place mainly at the level of strategies, which is formalized in the theory of strategy selection learning (SSL). In three studies we find that PCS-DM-L is more suitable to predict behavior than SSL.

1. Introduction

To make rational decisions and to avoid disastrous outcomes, individuals often need to infer criteria that are not directly observable. The distance of an approaching car, for example, cannot be directly observed but has to be judged from proximal depth-cues (e.g., texture-gradient). Similarly, whether a person lies or tells the truth in an interpersonal interaction cannot be directly observed but has to be inferred from cues that are probabilistically related to the distal criterion of interest. Correct inferences necessitate that a) individuals hold to some degree valid beliefs concerning the relation between cue and criterion and that they b) are able to use them in a ways that leads to reasonable outcomes.

While research in the tradition of the heuristics and biases program (e.g., Kahneman & Tversky, 1973) has focused on the issues of information usage and integration, the Brunswikian Lens Model (Brunswik, 1952) has inspired a wealth of research in which also the match between environmental structure and beliefs is taken into account. In this line of research, in which the structure of the environment and cue usage by individuals are investigated simultaneously, it has been shown that the accuracy of judgments in various domains is fairly high (judgmental achievement; see Karelaia & Hogarth, 2008, for a review). This finding can be considered to indicate that individuals hold reasonable accurate (implicit or explicit) beliefs concerning cue-criterion relations and are able to use them appropriately. Interestingly, however, studies investigating people's ability to learn cue criterion relations (e.g., Brehmer, 1969; Chasseigne, Mullet, & Stewart, 1997; Hammond, Summers, & Deane, 1973) yielded mixed results. It has been shown that judgmental accuracy can be increased by providing repeated feedback concerning the accuracy of a judgment (outcome feedback) in relatively simple tasks (e.g., Adelman, 1981; Doherty et al., 1988; Hirst & Luckett, 1992; Muchinsky & Dudycha, 1975) but less so in complex and uncertain tasks (e.g., Brehmer, 1980; Hoffman, Earle, & Slovic, 1981). Providing direct information about cue-criterion relations (task information feedback) is more efficient than outcome feedback (Balzer, Sulsky, Hammer, & Sumner, 1992; Reilly & Doherty, 1992; see also Brehmer, 1980) but from a practical perspective it has to be acknowledged that such explicit information will not be available in many real world situations.

One limitation of the work on the lens model was, however, that research followed an "as-if" approach that allowed to predict outcomes but remained largely silent concerning specifics of the underlying cognitive processes for information integration (Doherty & Brehmer, 1997; but see also Hoffmann, 1960). According to the work on bounded rationality (Simon, 1956), the development of psychologically reasonable models requires to take into account characteristics of the decision maker (e.g., limited memory capacity) and characteristics of the environment (e.g., local distribution of resources and information) simultaneously. One crucial aspect thereby is that people's capacity for deliberate cognitive calculations is limited and it is therefore assumed that "[p]eople satisfice—look for good-enough solutions—instead of hopelessly searching for the best" (Simon, 1990, p. 17).

Considering the task to infer criteria based on probabilistic cues, individuals already face a daunting task. Individuals (i) have to (explicitly or implicitly) know what the relevant cues are and how well they predict the criterion, they (ii) have to search for cues in memory or the environment, and they (iii) have to integrate them to determine the criterion and to select one of the available options. Importantly, in natural environments decision making does not end at this stage. People are involved in a continuous stream of inferences and actions in which feedback on (and consequences of) previous actions are used to reduce the gap between reality and predictions in order to improve the quality of subsequent choices (Clark, 2013). Specifically, to improve over time individuals (iv) have to learn. The question how learning in this continuous stream of inferences and choices can be cognitively modelled is investigated in the current paper. We contrast two fundamentally different mechanisms: individuals update cue-criterion relations only vs. individuals learn to select better strategies for integrating cues.

The currently prevailing view of adaptive decision making (e.g., Beach & Mitchell, 1978; Gigerenzer, Todd, & The ABC Research Group, 1999; Payne, Bettman, & Johnson, 1988, Scheibehenne, Rieskamp & Wagenmakers, 2013) assumes the second mechanism. According to this approach, individuals rely on a set of qualitatively different strategies for making decisions. Strategies differ in their complexity and most specified strategies are heuristics, that is, rules of thumb that simplify the decision-situation by ignoring parts of the information and using simple rules for information integration. It is assumed that individuals select strategies adaptively to environmental and situational demands and learn to adapt to the environmental structure by learning to select the most successful strategy. Learning to choose the right strategy for the right environment has been formalized in the Strategy Selection Learning Theory (SSL; Rieskamp & Otto, 2006), which has also received some empirical support (Rieskamp, 2006, 2008; Rieskamp & Otto, 2006).

Single-mechanism models for judgment and decision making, in contrast, assume that processes can be described by a single mechanism for information integration (e.g., Busemeyer, Pothos, Franco, & Trueblood, 2011; Busemeyer & Townsend, 1993; Dougherty, Gettys, & Odgen, 1999; Fiedler, 2000; Glöckner & Betsch, 2008; Lee & Cummins, 2004; Newell, 2005; Pleskac & Busemeyer, 2010; Thomas, Dougherty, Sprenger, & Harbison, 2008; Trueblood, Brown, & Heathcote, in press). Given that these models do not contain the concept of strategies in the first place, learning cannot take place on this conceptual level and has to be based on changes in cue-criterion relations only. In the current paper we will focus on one specific kind of models within the class of single mechanism models, namely Parallel Constraint Satisfaction (PCS) Models. PCS models are based on interactive activation networks that have been originally developed in the Parallel Distributed Processing framework for describing processes of perception (McClelland & Rumelhart, 1981) and beyond (McClelland, Rumelhart, & The PDP Research Group, 1986; Rumelhart, McClelland, & The PDP Research Group, 1986; see also McClelland, Botvinick, Noelle, Plaut, Rogers, Seidenberg, & Smith, 2010). PCS networks that model cognitive processes as coherence structuring between bottom-up (e.g., percepts, data, facts) and top-down (e.g., concepts, theories) influences, have been successfully applied to a wide variety of cognitive phenomena (e.g., Thagard, 1989; Holyoak & Thagard, 1989; Shultz & Lepper, 1996; Read, Vanman & Miller, 1997; Monroe & Read, 2008; Read, Monroe, Brownstein, Yang, Chopra, & Miller, 2010; Freeman & Ambady, 2011). Most importantly, PCS models for judgment and decision making (Glöckner & Betsch, 2008; Holyoak & Simon, 1999) were particularly successful in explaining multiple facets of behavior in probabilistic inference tasks in the domain of legal reasoning (e.g., Thagard, 2003; Simon, Chadwick & Read, 2004) and beyond (e.g., Brownstein, Read, & Simon, 2004; Glöckner, Betsch, & Schindler, 2011; Glöckner & Betsch, 2012; Glöckner & Bröder, 2011, 2014; Söllner, Bröder, Glöckner & Betsch, 2014).

One fundamental shortcoming of several PCS accounts developed so far is that they usually only account for information integration given a certain network structure. How this network structure emerges is usually beyond the scope of the model (Shulz & Lepper, 1996) and sometime rests on ad-hoc assumptions. Consequently learning and adaptation of networks is usually not part of the models and investigations reported so far focus on situations that do not resemble the full complexity of the stream of decisions introduced above. In this paper we aim to close this theoretical and empirical gap by extending the PCS model for decision making (Glöckner & Betsch, 2008) by a formalized learning-algorithm, testing its empirical adequacy to describe behavior.

1.1 The multi-strategy view: Learning to select strategies

The adaptive decision making perspective assumes that people rely on a collection of simple but successful algorithms that are selected contingent on environmental demands (e.g., Gigerenzer, Todd, & The ABC Research Group, 1999; see also Beach & Mitchell, 1978; Payne, Bettman & Johnson, 1988). The research program on the adaptive toolbox (Gigerenzer, Hertwig, & Pachur, 2011; Gigerenzer, Todd, & The ABC Research Group, 1999) is a prominent approach to the ecological rationality of decision making. Decision makers choose adaptively from a set of heuristics that are defined as simple, fast, and frugal. The heuristics proposed are postulated as cognitive algorithms of decision making. They are assumed to be ecologically rational and psychologically sound models of human decision making. For example, a decision maker following the prominent take-the-best heuristic (TTB) looks for cues in order of cue validities and decides in line with the first cue that discriminates between options (Gigerenzer & Goldstein, 1996). It has been shown that TTB performs comparable and sometimes even superior in making correct decisions in comparison to much more complex compensatory strategies in various decision environments (Czerlinski et al., 1999; Gigerenzer & Brighton, 2009; Martignon & Hoffrage, 1999) and under various environmental constraint (Hogarth & Karelaia, 2006a, 2006b; Katsikopoulos, Schooler, & Hertwig, 2010). It has also be shown in numerous studies that people behave in line with heuristics given environmental constraints like time pressure or high information-cost (e.g., Bröder, 2000a, 2000b, 2003; Todd et al., 2010). For the heuristics from the adaptive toolbox, a prominent learning algorithm has been proposed and tested: The strategy selection learning theory (Rieskamp, 2006, 2008; Rieskamp & Otto, 2006) assumes that people learn to pick the most successful strategy by experience through external feedback. SSL is not restricted to the simple heuristics from the "adaptive toolbox". The main assumption is that through feedback, reinforcement commences at the level of strategies, meaning a set of cognitive operations employed to solve a task.

1.2 Single mechanism models: Learning to adapt cue-weights in network models

The second complementary perspective focuses less on limitations of deliberate cognitive capacity, drawing instead on the broad knowledge concerning the existence of computationally powerful cognitive mechanisms (see Glöckner & Witteman, 2010, for a review). It has been repeatedly shown that cognitive mechanisms of memory activation and perception (e.g., McClelland & Rumelhart, 1981), the construction of meaning (Thagard, 1988, 1989), attitude formation (e.g., Monroe & Read, 2008), and many more can be modeled as PCS networks. These models have also been applied to account for judgment and decision making processes (e.g., Betsch & Glöckner, 2010; Glöckner & Betsch, 2008; Holyoak & Simon, 1999; Thagard & Millgram, 1995). In the most recent implementation of PCS for decision making (Glöckner & Betsch, 2008; Glöckner, Hilbig, & Jekel, 2014), it is assumed that the default mode of cue integration is automatic and parallel. Deliberate processes only intervene if the parallel processes of cue integration do not lead to a sufficiently coherent interpretation. PCS for probabilistic decision making has been shown to lead to good decisions in various environments (Jekel, Glöckner, Fiedler, & Bröder, 2012) and to be a valid account of decision-behavior in various domains (Glöckner & Betsch, 2008, 2012; Glöckner, Betsch, & Schindler, 2010; Glöckner, Heinen, Johnson, & Raab, 2010).

For PCS approaches to decision making, work on adaptation to the environment by learning from experience is mathematically complex, and convincing solutions are still largely missing. Additionally, the neglect of learning models for PCS networks has been criticized (Smith, 1996; Van Overwalle, 1998). However, there are some noteworthy attempts to implement learning rules into related network models, which can be used as starting point (Johnson, Zhang, & Wang, 1997; Vanhoomissen & Van Overwalle, 2010; Van Overwalle & Siebler, 2005). In the literature on network learning, the Delta-rule—an algorithm of backpropagation that shares properties with the Rescorla-Wagner model of learning (Rescorla & Wagner, 1972; Van Overwalle, 2007)—has proven to be successful (Bechtel & Abrahamsen, 2007; Rumelhart, Hinton, & Williams, 1986; Rumelhart & McClelland, 1986). In the following, we introduce the formalization of the parallel-constraint satisfaction network of modeling learning in probabilistic decision making. In contrast to the SSL theory, this approach assumes that validities for the cues are updated by feedback, thus changing weights in the network structure. The integration mechanism for the information rests on spreading activation and does not change through learning.

2. Implementing learning in PCS

A model of learning in probabilistic decision-making needs to address how people learn probabilistic cue-criterion relations between decision-trials based on feedback, how people make decisions in each trial based on those learned relations, and how both processes relate to each other. The process of decision making in PCS has already been formalized elsewhere (Glöckner & Betsch, 2008) and will be briefly summarized in the following section. We then introduce the extension of PCS by adding the modified Delta-rule for learning of cue-criterion relations between trials. To facilitate understanding of the formalization of PCS introduced next, we also provide the code of the implementation of PCS in the statistical software-package R (R Team, 2014) in Appendix C.

2.1 Primer of PCS

The decision-situation—that is, participants' representation of the cue pattern, subjective validities for cues, and the decision-alternatives—is represented in a symbolic network-model (Figure 1). The decision-process is simulated as an iterative process of spreading activation in the network.

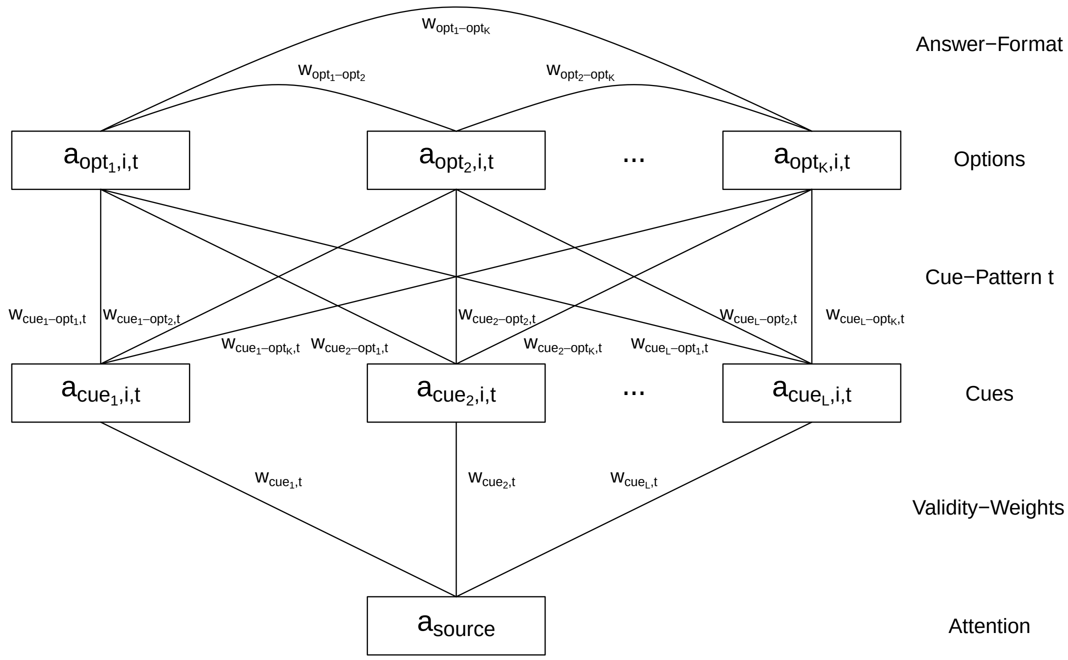

Figure 1. The parallel-constraint-satisfaction network-model of decision making (c.f. Figure 1, Glöckner & Betsch, 2008, p. 2).

2.1.1 Decision-situation as represented in the network-model

The network-model consists of nodes that are interconnected. Cues are represented in the middle layer of the network as nodes $\{cue_1, cue_2, \ldots, cue_L\}$, choice-options are represented in the upper layer as nodes $\{opt_1, opt_2, \ldots, opt_K\}$. From a general source node constant activation spreads into the network by an iterative updating-algorithm that simulates the decision-process. In total there are $N = L + K + 1$ nodes in the model in three layers. The cue pattern of trial t is represented as weights

$\{w_{cue_1-opt_1,t},\ldots,w_{cue_L-opt_K,t}\}$ attached to the bi-directional connections between $K = 2$ options and $L = 5$ (Exp. 1 and 2), and $L = 6$ (Exp. 3) cues. Weights can be excitatory (positive) or inhibitory (negative): A weight between a cue and an option receives a value of $+.01$ when it speaks for the option and a value of $-.01$ when it speaks against the option. The subjective validities of the cues are represented as weights $\{w_{cue_1,t},\ldots,w_{cue_L,t}\}$ attached to the connections between the source-node and the cue-nodes. Validity-weights are positive when the presence of a cue for an option is associated with a higher likelihood of being the better option and negative when the absence is associated with a higher likelihood. All option-nodes are connected with negative weights $\{w_{opt_1-opt_2,t},\ldots,w_{opt_{(K-1)}-opt_K,t}\}$. In the simulations, we set $w_{opt_1-opt_2,t}$ equal to $-.2$, mapping the answer-format of a forced-choice between two options in the decision task (i.e., choosing an option inhibits choosing the alternative option).[1] How the process of decision-making is simulated in the network in trial t is explained next.

2.1.2 Simulating the process of decision-making

Nodes can change their activation a in an iterative process $i = \{1, 2, \ldots, I\}$. In the first iteration $i = 1$, the source node is activated with an activation of 1. In the second iteration $i = 2$, the attention is directed towards the cues: Weighted activation spreads from the source-node to the cue-nodes. In the third iteration, activation continues to spread from the source-node to the cue-nodes but also spreads from the cue-nodes to the option-nodes. The property of bi-directional links distinguishes PCS from other prominent network models (Gluck & Bower, 1988) and leads to distinct predictions concerning coherence effects in the process of decision making (Glöckner, Betsch, & Schindler, 2010; Holyoak & Simon, 1999). In the fourth iteration, activation spreads from the source-node to the cue-nodes to the option-nodes but also from the option-nodes back to the cue-nodes. The input of activation for each node p for iteration i and trial t is calculated by (cf., Glöckner & Betsch, 2008, Formula 2, p. 218):

$$input_{node_p,i,t} = \sum_{q=1,q\neq p}^{Q=N} w_{node_q-node_p,t} \times a_{node_q,i,t}. \tag{1}$$

The input for each node results from the weighted activations of the nodes attached to the target-node. The activation a for each node p at iteration $i+1$ results from the previous activation at t and a weighted input to the node according to the following function (cf., Glöckner & Betsch, 2008, Formula

[1] Conclusions are not dependent on the exact value used.

1, p. 218):

$$a_{node_p,i+1,t} = a_{node_p,i,t} \times (1 - decay) + input_{node_p,i,t} \times \begin{cases} (a_{node_p,i,t} - floor) & \text{if } input_{node_p,i,t} < 0 \\ (ceiling - a_{node_p,i,t}) & \text{if } input_{node_p,i,t} \geq 0 \end{cases} \tag{2}$$

Activation at iteration $i+1$ results from the summed activation at i weighted by the factor $decay = .1$ (i.e., the impact of prior activations decrease over iterations) and the weighted input (i.e., activations can only range between $ceiling = 1$ and $floor = -1$ in a sigmoid function). The overall negative energy in the network is defined by (c.f., Read, Vanman, & Miller, 1997, p. 30):

$$energy_{i,t} = -\sum_{p=1}^{P=N} \sum_{q=1,q\neq p}^{Q=N} w_{node_p-node_q,i,t} \times a_{node_p,i,t} \times a_{node_q,i,t}. \tag{3}$$

The activation of connected nodes are multiplied with the weight of the connection. Negative energy approximates the minimum over iterations of updating activations given the constraints of the network (Hopfield, 1982, 1984; Read, Vanman, & Miller, 1997). When the negative energy from iteration $i = I - 1$ to iteration $i = I$ changes negligibly, updating of node-activations is stopped and the simulation of the decision-process in the network is terminated. The option-node with the highest activation at iteration I is the predicted option chosen by the participant for trial t. The number of iterations I is used as a proxy for predicting decision-times. Participants' learning of cue-validities is represented by updating validity-weights after receiving feedback in trial t according to the modified Delta-rule as explained next.

2.2 Simulating learning: A modified Delta-Rule

Validity-weights in the network are updated using a modified Delta-rule based on the final activations of the option-nodes. The difference between the observed activations of option nodes $a_{opt_k,I,t}$ and a desired level of activation $d_{a_{opt_k,I,t}}$ determines the weights $w_{cue_l,t}$ for time $t + 1$. Validity-weights are changed by $\Delta w_{cue_q,t}$ according to the following modified Delta-rule (cf. Bechtel & Abrahamsen, 2007; Rumelhart & McClelland, 1986):

$$\Delta w_{cue_q,t} = \lambda \times \sum_{p=1}^{P} [(d_{a_{opt_p,I,t}} - a_{opt_p,I,t}) \times w_{cue_q-opt_p,t}] \tag{4}$$

A larger difference between desired and observed activation $d_{a_{O_k}} - a_{O_k}$ for trial t at the final iteration I of the decision process results in a larger change in weights. The learning rate λ moderates the impact of a single decision trial on the change of the weights and thus moderates the speed of learning. For example, in the decision trial $t = 1$ the weight $w_{cue_1,t}$ of $cue_{1,t}$ (i.e., its subjective validity) might be positive and the weight $w_{cue_1-opt_1,t}$ negative, that is, cue 1 speaks against option 1. After the network settled its activations in the final iteration I, the activation for option 1 $a_{opt_1,I,t}$ might be positive due to the other cues $2 \ldots I$ in the network. Let us assume the desired activation $d_{a_{opt_{1,I,t}}}$ was positive at .6, that is, PCS chooses the correct option ($a_{opt_1,I,t}$ and $d_{a_{opt_{1,I,t}}}$ are both positive) irrespective of the incorrect prediction of cue 1 for option 2.[2] Updating via the Delta-rule results in a lower weight $w_{cue_1,t}$ since a lower (i.e., ultimately negative weight) would have produced a correct prediction for option 1.[3] Note that λ is the only free parameter in the equation, the desired activation for option-nodes and the observed activation of option-nodes follow from the PCS simulation in the preceding decision. Updating of weights is done according to:

$$w_{cue_p,t+1} = w_{cue_p,t} + f(\Delta w_{cue_p,t}). \tag{5}$$

The function $f(\Delta w_{cue_it})$ leads to a sigmoid function to assure that net-weights can only vary between $+1$ and -1 (McClelland & Rumelhart, 1981):

$$f(\Delta w_{cue_p,t}) = \Delta w_{cue_p,t} \times \begin{cases} (1 - w_{cue_p,t}) & \text{if } w_{cue_p,t} \geq 0 \\ (1 + w_{cue_p,t}) & \text{if } w_{cue_p,t} < 0 \end{cases} \tag{6}$$

2.3 Deriving choice probabilities

We test two different algorithms to derive choice-probabilities from the activation of option-nodes in PCS that tap different psychologically plausible properties of the decision maker.

[2]The desired activation is not set to -1 and +1 because the proposed network with five or six cues and two options can only maximally and minimally (i.e., all validity-weights set at 1 and all cues pointing towards one option) produce activations for option-nodes between $\simeq -.63$ and $\simeq .63$.

[3]For matters of simplification in the example, only $d_{a_{opt_{1,I,t}}}$ and $a_{opt_1,I,t}$ are compared. To determine $\Delta w_{cue_1,t}$, it is further necessary to make the remaining comparisons $d_{a_{opt_{2...K,I,t}}}$ with $a_{opt_{2...K,I,t}}$ for $w_{cue_{1,t}-opt_{2...K,t}}$ (see sigma sign in the formula).

2.3.1 Sensitivity to the strength of evidence

The first approach relies on the difference in the activations of the two option nodes in PCS as a proxy for choice-probabilities. A high activation for one option and a low activation for the the other option result from an interplay between an unambiguous cue-pattern in a trial (e.g., most cues favor one option) and high net-weights for connections between the source node and validity nodes (i.e., a clear cue-pattern does not necessarily maximize the difference between node-activations when net-weights for validities are low or indistinct from each other). As indirect evidence that differences in node-acitvations map to choice probabilities, it has been shown in various studies that those differences relate to observed confidence judgments in participants' decisions (Glöckner & Betsch, 2008). Thus, we assume that an increasing difference in node-activations for options is related to more extreme predicted choice-probabilities. Thus, for trial t given activation $a_{opt_c,I,t}$ chosen by the participant and the activation for the option not chosen $a_{opt_{nc},I,t}$, the choice probability can be calculated by (cf. Glöckner, Heinen, Johnson, & Raab, 2012, p. 331):

$$p_t = \frac{e^{\gamma \times a_{opt_c,I,t}}}{e^{\gamma \times a_{opt_c,I,t}} + e^{\gamma \times a_{opt_{nc},I,t}}}. \tag{7}$$

For each cue-pattern $t = \{1 \ldots T\}$, the exponential of the activation of the option chosen by the participant $a_{opt_c,I,t}$ is divided by the sum of the exponentials of the activations for both options. The probability of the chosen option is high when the activation of the node representing the chosen option is high in comparison to the unchosen option. Note that the parameter γ accounts for individual differences in how differences in node-activations map to choice-probabilities. A decreasing γ results in a decreasing sensitivity to the activations of the options in the network (e.g., for $\gamma = 0$, probabilities for both options are .5, that is, a participant is insensitive to differences in node activations). The maximum log-likelihood of all choices of a participant can be calculated by finding for each participant the individual parameters of the learning rate λ in the interval $[0, 2]$ and the sensitivity in the transformation function γ in the interval $[0, 5]$ that maximize:

$$ln(L_{\text{PCS}_{\text{Transf}}}) = \sum_{t=1}^{T} ln(p_t). \tag{8}$$

2.3.2 Noisy learning

The second approach assumes that learning of cue-validities and thus updating of validity-weights in PCS is not deterministic (Equation (5)) but noisy. This approach shares the same Delta-learning-rule for updating net-weights with the other approach of implementing learning in PCS but differs on how deterministic predictions of PCS are transformed into probabilistic predictions by assuming that updating of cue-validities is partially probabilistic due to unsystematic error in the learning process. To simulate noisy updating of validity-weights in PCS for trial t, a random number from a normal distribution with a mean of 0 and a standard deviation of ρ is added to the validity-weight after each trial t:

$$w^{\rho}_{cue_p,t+1} = w_{cue_p,t+1} + N_t(0,\rho). \tag{9}$$

To determine the prediction for the choice-probability for an option in trial t for a participant, we simulate $N = 1{,}000$ PCS-models solving all T cue-patterns differing only in the random-noise component $N_t(0,\rho)$ with a standard deviation ρ fixed at a specific value. We then use the percentage of PCS-models that choose the option selected by a participant in the study as the predicted choice-probability for the option (e.g., when 900 out of $1{,}000$ simulated participants choose option 1 for cue-pattern t, $p(opt_{1,I,t}) = .90$). The maximum log-likelihood of all choices of a participant therefore results from the individual parameter of the learning rate λ in the interval $[0,2]$ and the standard deviation of the error ρ in the interval $[0,.05]$ that maximize:

$$ln(L_{\mathrm{PCS_{Noise}}}) = \sum_{t=1}^{T} ln[p(opt_{c,I,t})]. \tag{10}$$

From a modelling perspective, this approach also circumvents the need for addtional assumptions on how to model error (e.g., trembling-hand error in SSL; error in accordance with a logistic-function in $\mathrm{PCS_{Transf}}$) when the prediction of a deterministic model does not overlap with an observed choice by a participant (Myung, Pitt, & Kim, 2005).

3. Competitor model: Strategy-Selection-Learning theory

The strategy selection learning theory from the adaptive toolbox has been introduced elsewhere in detail (Rieskamp, 2006, 2008; Rieskamp & Otto, 2006); we limit our description of the theory to a conceptual summary and spell out the formalization in Appendix A. People possess a set of strategies that they can pick from to solve a decision task. A prominent non-compensatory strategy is the Take-the-best heuristic: A decision is based on the most valid cue that discriminates between choices options. A prominent compensatory strategy is Weighted Additive: Cues are weighted by validities and then summed for each option. The option with the higher weigthed sum is predicted to be chosen by the participant (Appendix A.2). This set of strategies is not exhaustive but the authors of SSL (s.a.) included those two strategies representative for (non-)compensatory strategies only in prior studies and we therefore also do so.

According to this strand of theorizing, people need to update cue-validities between trials because strategies rely on cue-validities (i.e., Take-the-best: cue-search by validities; Weighted Additive: Weighting by validities) and people need to learn which of the two strategies to pick in order to adapt to a (non-)compensatory environment. Validities are updated according to an event-counter (Gigerenzer & Goldstein, 1996; Gigerenzer, Tood, & The ABC Research Group, 1999; Martignon & Hoffrage, 1999): The validity is the number of trials in which the cue signaled the better option divided by the number of all trials in which the cue signaled one of the options (Equation (A.1), Appendix A). Although the psychological plausibility of an event-counter has been discussed (Dougherty & Franco-Watkins, 2008; Gigerenzer, Hoffrage, & Goldstein, 2008), research shows that people are able to process frequency information with ease (Gigerenzer & Hoffrage, 1995) and represent frequencies of events very accurately (Fiedler, 2008; Zacks & Hasher, 2002). How people learn to choose the best strategy between trials is formulated in SSL. SSL assumes a reinforcement-mechanism to take place: A strategy that performed well in a trial is more likely used in the next trial. Participants are assumed to enter the lab with a prior expectancy that a strategy performs well based on (e.g.) prior experience with the strategy outside the lab. A participant with a high expectancy for (e.g.) Weighted Additive will likely choose this strategy in the first trial. In case the participant chooses to apply Weighted Additive, she might fail to apply this strategy correctly with a certain trembling-hand error-probability (i.e., unsystematic noise). Finally, when the application of the strategy leads to a correct decision, Weighted Additive is more likely used in the next trial (and therefore Take-the-best is less likely used) although the impact of a single learning-experience may differ between participants as also captured by the model.

4. Study 1: Learning in dynamic environments

In the first study we test the accuracy of our extension of learning in PCS against SSL and single-strategy models. We use a dynamic environment with changing probabilistic relations between cues and criterion in the course of the experiment.

4.1 Method

4.1.1 Participants

75 participants (42 female, mean age = 24 years) were recruited from the Decision Lab Subject Pool of the Max-Planck-Institute using the online recruiting tool ORSEE (Greiner, 2004). Participants received a show-up fee of 5 Euros (5.5 Dollars) and additional performance contingent payment of maximally 7.76 Euros (8.5 Dollars).

4.1.2 Materials and Design

We used a hypothetical stock-market game adopted from previous research (Bröder, 2003). Participants were asked to select the more profitable of two stocks in a series of trials consisting each of a new pair of stocks (Figure 2). Participants were provided with information from five experts (i.e., cues) that made recommendation concerning whether the expected profitability of the respective stock was good or bad (i.e., binary cues). The recommendations of each expert could differentiate between the two options or could be indifferent between options (i.e., both options recommended).

The validities of the experts were varied as between-subjects factor with $val_{cond_1} = \{.90, .80, .70, .65, .60\}$, $val_{cond_2} = \{.95, .65, .63, .61, .59\}$, and $val_{cond_3} = \{.73, .71, .69, .67, .65\}$. Participants received information on the ranking of the experts according to their validity in the first round (i.e., expert A was the most valid cue, expert B the second most valid and so on). We generated all 121 unique cue-patterns with five binary cues (Jekel, Glöckner, & Fiedler, 2010). We defined the better of two option for each trial as the more probable option in accordance with (naïve) Bayes given the cue-pattern and the validities of the experts in the environment. That is, for each cue-pattern we calculated for each option the posterior-probability given the validities of the environment and defined the option with the higher posterior probability to be the better option (Equation (A.3)). After each decision, participants received feedback about the better option and could thereby update subjective validities for the five

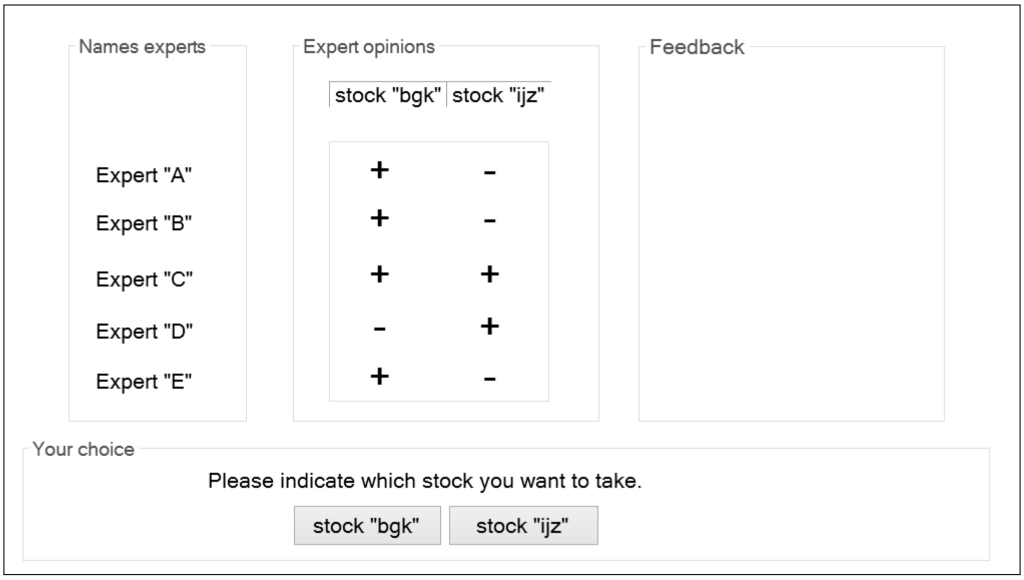

Figure 2. Screenshot of a trial in the first study (translated from German). The binary cue pattern (support indicated by pluses and rejection by minuses) for two stocks is displayed. After a decision for one of the stocks, the participant receives feedback about the better option in the right window of the display.

experts and/or the success of decision strategies. Participants received a payment of 2 Cents for a decision for the better stock. Cue-patterns were repeated in three blocks with random order in each block. After each block of 121 decisions, participants were warned that the following 121 trials are from a different stock-market and the five experts may change in the validity of predicting the better stock. The most valid cue in each condition decreased close to chance-level with $val_{cue_1} = .51$ in the final round only.

4.2 Procedure

Participants received instructions on the stock-market game. After completing all 363 trials (3 rounds $\times$ 121 cue-patterns), participants were asked to order experts from the most valid to the least valid for the last round in which the most valid cue changed to the least valid cue. For a correct ordering, participants received an additional 0.50 Euros (0.55 Dollar). Finally, participants were asked for demographic data, debriefed, and paid.

4.3 Models

We compared the accuracy to predict participants' choices for six models that are briefly summarized in Table 1: The two network implementations of PCS-DM-L introduced above, the Strategy Selection Learning Theory (SSL), and single strategies from the adaptive toolbox.

Table 1. Short description of all six models from two model classes (i.e., network models and adaptive toolbox) compared in all three studies.

		Network models		
#	Information integration	Abbrev.	Free parameters	Description
1	Parallel Constraint Satisfaction network-model with modified Delta-Learning-Rule and transformation function	$\text{PCS}_{\text{Transf}}$	learning rate λ, exponential choice-rule with sensitivity γ	Pick the option that is maximally-coherent to the pattern of evidence and the quality of evidence as learned with a rate λ from past trials according to the Delta-rule. The probability of choosing the option is dependent on the sensitivity γ to differences in evidence for options.
2	PCS with noisy learning	$\text{PCS}_{\text{Noise}}$	learning rate λ, noise ρ	The probability of choosing the option is dependent on the extent of random-noise from a normal distribution with standard deviation ρ when updating validities.
		Adaptive toolbox		
3	Strategy Selection Learning Theory	SSL	initial preference β_{TTB}, strength of initial preference w (i.e., learning rate $1/w$), strategy application-error ϵ	Pick the option according to TTB or WADD with probability $1 - \epsilon$ in proportion to the strategies' expected success resulting from an initial preference for strategies β_{TTB} and $\beta_{\text{WADD}} = 1 - \beta_{\text{TTB}}$ with strength w and the experienced success of each strategy from past trials.
		Single strategies from the adaptive toolbox		
4	Take-the-best	TTB	strategy application-error ϵ_T	Decide in line with the most-valid cue that discriminates between options with probability $1 - \epsilon_T$.
5	Weighted Additive	WADD	strategy application-error ϵ_W	Pick the option with the highest sum of cues weighted by validities with probability $1 - \epsilon_W$.
6	Naïve Bayes	RAT	strategy application-error ϵ_R	Pick the option with the highest odds-ratio with probability $1 - \epsilon_R$.

4.4 Results

4.4.1 Performance

Participants choose the better option in 92% of all trials in the first round (Figure 3, left). The mean performance does not change significantly in the second round and drops to 83% in the third round when the most valid cue changes to the least valid cue. The cumulative performance shows that learning takes place in each round (Figure 4, left three lines) and that participants quickly adapt to the environment (i.e., around 90% correct decisions after 31 trials in the first round). A multilevel-regression model with performance per round as criterion and round and environment (i.e., cue-validities for experts) as predictors with dummy coding allowing for a random intercept and random slope for each round conditioned on participants shows that performance is significantly lower in the third round in comparison to the first round ($b = -0.08$, $t(148) = -14.48$, $p < .001$) and participants perform significantly higher in the third environment with $val_{cond_3} = \{.73, .71, .69, .67, .65\}$ than the first environment with $val_{cond_1} = \{.90, .80, .70, .65, .60\}$ ($b = 0.02$, $t(72) = 2.45$, $p < .05$). All other predictors are insignificant (i.e., no significant differences between round 1 and round 2, no significant differences between environment 1 and 2). Participants were also asked to order experts according to their validity in the final round. Although the most frequent answer is the correct ordering (i.e., fifth cue least valid due to change), only 8 out of 75 participants (i.e., $\sim 11\%$) give this answer followed by the next most frequent category of five participants sticking to the order of experts in the first two rounds. Thus, a majority of participants (i.e., $\sim 89\%$) err more or less when ordering cues by validity.

4.4.2 Model parameters and model comparison

The mean of individually fitted model-parameters for SSL (Table 2, upper part) are comparable to prior studies (Rieskamp & Otto, 2006, Table 2, p. 220) except for the application error ϵ being relatively higher for SSL and also for the single strategies which might result from the dynamic change in the environment. The learning rate λ differs between the two versions of PCS which is due to the different implementations of deriving choice probabilities.

PCS_{Transf} predicts 90% of all participants' choices correctly, PCS_{Noise} and SSL perform slightly worse with 89% and 88% of all choices correct (Table 3, upper part). Based on the posterior probabilities for each participant and model (Appendix B), a majority of 89% of all participants can be best explained

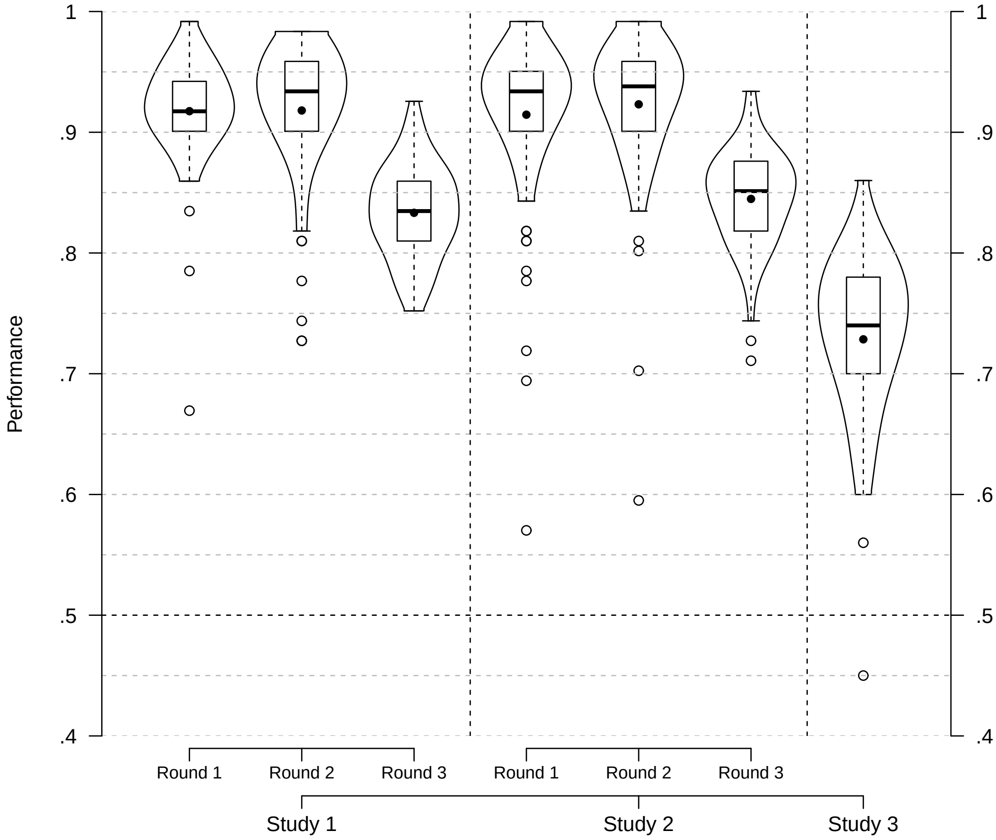

Figure 3. Overlap of participants' choices with the (naïve) Bayesian solution (i.e., performance) for round 1 to 3 in study 1 (left), study 2 (middle), and for study 3 (right). Violin plots are displayed: Means are black dots, medians are black thick lines, the borders of the box indicate the lower or upper quartile, whiskers indicate the minimum or maximum data point within the range of $1.5\times$ the difference between quartiles subtracted or added from the lower or upper quartile (given there are no outliers), white dots indicate outliers (i.e., data outside the range of whiskers), and shapes around the box-plots indicate the density distribution of the data.

by $\text{PCS}_{\text{Noise}}$, a minority of 9% can be best explained by $\text{PCS}_{\text{Transf}}$, and only 1% of participants can be best explained by SSL. Posterior probabilities for participants classified is close to 1 for all models (last column, Table 3). The overall likelihood for $\text{PCS}_{\text{Noise}}$ (i.e., across all participants) is overwhelming for $\text{PCS}_{\text{Noise}}$ with $\text{PCS}_{\text{Noise}}$ being e^{1061} more likely than $\text{PCS}_{\text{Transf}}$ and e^{1977} more likely than SSL. Single strategies perform on all measures poorly in comparison to both PCS-implementation of learning and SSL.

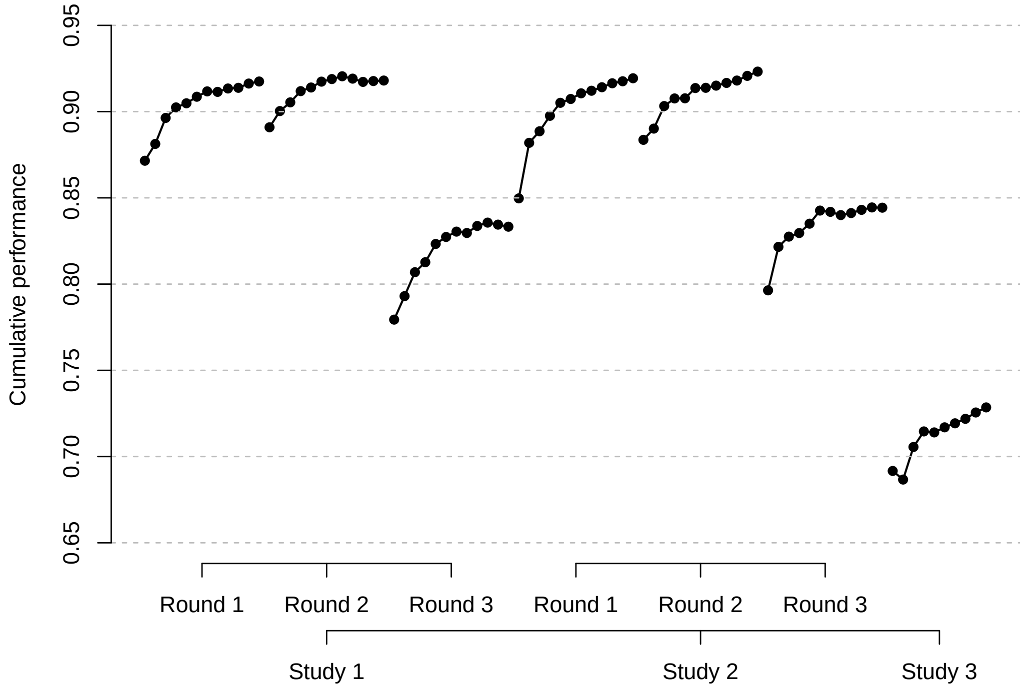

Figure 4. Cumulative participants' performance for round 1 to 3 from 11 to 121 trials (study 1 and 2) and from 10 to 100 trials (study 3) in steps of 10 trials.

Table 2. Mean and standard errors for parameters for all models for study 1 to 3.

	PCS_{Transf}		PCS_{Noise}		SSL			TTB	WADD	RAT
	λ	γ	λ	ρ	w	β_{TTB}	ϵ	ϵ_{TTB}	ϵ_{WADD}	ϵ_{RAT}
S 1										
Mean	0.63	2.17	1.4	0.011	54.26	.31	.08	.19	.13	.13
SE	0.06	0.04	0.05	0.001	5.44	.02	<.001	.01	<.001	<.001
S 2										
Mean	0.65	2.22	1.4	0.011	31.16	.31	.09	.20	.13	.12
SE	0.05	0.04	0.04	0.001	4.31	.02	<.001	<.001	<.001	<.001
S 3										
Mean	0.48	1.17	0.84	0.018	36.64	.39	.22	.39	.34	.36
SE	0.05	0.04	0.07	0.001	6.08	.04	.01	.01	.01	.01

4.5 Summary

Participants are able to adapt to the structure of the environment: Performance is already high with 92% of all choices correct in the first round of 121 decision trials. Although performance drops when the

Table 3. Percentage of participants best explained, choices correctly predicted, log-Likelihoods and Bayesian information criterion (BIC) summed over participants, and posterior probabilities averaged over participants for each model, each study, and for all studies.

	% Partic	% Choices	Sum log(Lik)	Sum BIC	Mean Posterior
S 1, N = 75, C = 27,225					
PCS_{Transf}	9	90	-8344	17571	.95
PCS_{Noise}	89	89	-7283	15449	$\simeq 1$
SSL	1	88	-9038	19402	$\simeq 1$
TTB	0	81	-12947	26336	—
WADD	0	87	-10263	20967	—
RAT	0	87	-10375	21193	—
S 2, N = 100, C = 36,300					
PCS_{Transf}	13	91	-10864	22907	.87
PCS_{Noise}	85	89	-9814	20807	.99
SSL	0	88	-12038	25843	—
TTB	1	80	-17673	35936	.90
WADD	1	87	-13607	27803	.54
RAT	0	87	-13351	27291	—
S 3, N = 60, C = 6,000					
PCS_{Transf}	57	77	-3181	6914	.89
PCS_{Noise}	27	72	-3352	7256	.86
SSL	5	70	-3461	7752	.65
TTB	7	61	-3921	8118	.77
WADD	0	66	-3747	7770	—
RAT	5	64	-3853	7983	.70
Overall, N = 235, C = 69,525					
PCS_{Transf}	23	89	-22388	47392	.89
PCS_{Noise}	71	88	-20449	43513	.98
SSL	2	86	-24537	52997	.74
TTB	2	79	-34541	70389	.79
WADD	0	85	-27616	56540	.54
RAT	1	85	-27579	56467	.70

Note. N = Number of participants, C = Number of choices; Percentages may not add up to exactly 100% due to rounding.

structure of the environment changes, participants are able to adapt and their performance is high with 83% correct decisions. Participants' adaptive learning in dynamic environments can be best explained by PCS with a modified Delta-Rule and noisy learning. The evidence for PCS is overwhelming in comparison to the models from the adaptive toolbox.

5. Study 2: Learning in dynamic environments with positive or negative consequences

In the following study, we aimed to generalize and test for the stability of results. In the first study, participants received a bonus for a correct decision while there was no negative outcome for an incorrect decision. In many decisions in the real-world, an incorrect decision can lead to a negative outcome. A comprehensive theory of probabilistic learning should account for learning by reinforcing correct decisions and learning by punishing incorrect decisions (i.e., avoidance learning). We therefore tested the effect of different consequences for decisions (i.e., receiving a bonus for a correct decision or being punished for an incorrect decision) on the ability of the models to explain choices. We also used an independent pool of participants to test for the stability of the effects in different subject-populations.

5.1 Methods

5.1.1 Participants

100 participants (66 female, mean age = 21 years) were recruited from the University of Mannheim. Participants received performance contingent payment of maximally 9.28 Euros (10.29 Dollars).

5.1.2 Materials, Design, and Procedure

We used the same stock-market game and the same cue-patterns as in the first study. In difference to the first study, we only used the environment val_{cond_2} and val_{cond_3} as typical non-/compensatory environments as between-subjects factor. Participants were also informed about the ranking of the experts according to their validity in the first round. Additionally, half of the participants were rewarded 0.025 Euros for a correct decision and the other half was punished by losing 0.025 Euros from an initial payment of $0.025 \times 363 = 9.075$ Euros for an incorrect decision. The procedure was also identical to the first study. All 121 cue-patterns were repeated in three rounds with random order in each round. Similar to the first study, the validity of the most valid expert changed close to chance-level with $val_{cue_1} = .51$ in the final round. For a correct ordering of experts according to their validity in the final round, participants received an additional payment of 0.20 Euros (0.22 Dollars).

5.2 Results

5.2.1 Performance

Results from the first study were replicated. Participants already choose the better option in 91% in the first round (Figure 3, middle). The mean performance slightly increases in the second round to 92% and drops to 84% in the third round. Similar to the first study, we ran a multilevel-regression model with performance per round as criterion and round, environment (i.e., cue-validities for experts), and type of payoff (i.e., reward versus punishment) as predictors with dummy-coding (i.e., round 1, environment 2, and punishment as control) allowing for a random intercept and random slope for each round conditioned on the participant. Results are similar to the first study. Participants perform significantly worse in the third round in comparison to the first round ($b = -0.07$, $t(198) = -10.21$, $p < .001$), and perform significantly better in the third environment in comparison to the second environment ($b = 0.019$, $t(97) = 2.96$, $p < .01$). Participants receiving a reward for a correct decision versus a punishment for an incorrect decision perform slightly worse with $-.008$ less correct choices (i.e., 2.90 choices on average) in comparison to the condition with rewards but this difference is statistically insignificant ($t(97) = -1.27$, $p = .21$). Similar to the first study, the number of participates who incorrectly stick to the order of cues in the first two rounds and the number of participants correctly indicating the order is roughly the same with a ratio of 9 to 8, a majority of 83 people gave other incorrect orderings of cues. Thus, a majority of participants in study 2 was also unable to order experts according to validity correctly.

5.2.2 Model parameters and model comparison

Mean model-parameters fitted per participant (Table 2, second section) are almost identical to the first study. $\text{PCS}_{\text{Transf}}$ predicts 91% of all participants' choices correctly, $\text{PCS}_{\text{Noise}}$ and SSL perform slightly worse with 89% and 88% of all choices (Table 3, middle part). A majority of 85% of all participants can be best explained by $\text{PCS}_{\text{Noise}}$, a minority of 13% can be best explained by $\text{PCS}_{\text{Transf}}$, and none of the participants can be best explained by SSL. The posterior probabilities for the classified participants is again high (last column, Table 3). The overall likelihood for $\text{PCS}_{\text{Noise}}$ is overwhelming for $\text{PCS}_{\text{Noise}}$ with $\text{PCS}_{\text{Noise}}$ being e^{1050} more likely than $\text{PCS}_{\text{Transf}}$ and e^{2518} more likely than SSL. Similar to the first study, single strategies perform on all measures poorly in comparison to both PCS-implementation of learning and SSL.

5.3 Summary

Participants already adapt to the environment in the first learning-round with 91% correct choices and are also able to adapt to a change in the environment with still 84% correct choices in the third round. PCS with a modified Delta-Rule predicts participants' choices best in dynamic environments for different types of consequences (i.e., reward versus punishment).

6. Study 3: Increasing the complexity of the decision-task

In the two preceding study, participants already performed at 90% correct choices after the first round of 121 trials. We therefore increased the complexity of the learning task in the third study by increasing the number of cues from five to six and by increasing the difficulty of decision-trials to test whether the PCS-model can also account for decision making in more complex decision-environments.

6.1 Methods

6.1.1 Participants

60 participants (34 female, mean age = 24) were recruited from the Decision Lab Subject Pool of the Max-Planck-Institute using the online recruiting tool ORSEE (Greiner, 2004). The study was run with another unrelated study following, participants received a performance contingent payment of maximally 4 Euros (4.43 Dollars).

6.1.2 Materials, Design, and Procedure

We used the same stock-market game (Figure 2). Contrary to the two preceding studies, we increased the complexity of the task by increasing the number of experts from five to six. The validity of the experts (cues) was set to $val = \{.89, .77, .72, .65, .60, .55\}$. Identical to the first two studies, participants were informed about the ranking of the experts. We also manipulated if participants received 0.04 Euros for a correct decision or lost 0.04 Euros for an incorrect decision. Contrary to the preceding studies, participants played a single round of 100 cue-patterns and the validity of the cues did not change. A set of difficult cue-patterns was selected by choosing from the set of all possible unique cue-patterns 80 patterns that resulted in the lowest log-Odds for the better option according to naïve

Bayes (i.e., patterns for which the evidence for both options is similar). We picked another 20 random cue-patterns for the complete set of 100 trials. After completing all 100 trials, participants were asked for demographics, debriefed, and paid.

6.2 Results

6.2.1 Performance

Participants' mean performance was 73 out of 100 correct cue-patterns (Figure 3, right) and therefore lower as observed in the preceding studies. Thus, our manipulation of increasing complexity by selecting only difficult tasks and increasing the number of cues was successful. Participants who receive rewards for correct decisions versus punishment perform with .71 versus .74 of correct choices slightly worse although this difference does not reach conventional levels of significance ($t(58) = -1.58$, $p = .06$).

6.2.2 Model parameters and model comparison

The learning rate γ is with 0.48 and 0.84 lower for both PCS-implementations (Table 2, lower part) in comparison to the preceding studies. The exponent γ is also lower with 1.17 versus 2.17 for $\text{PCS}_{\text{Transf}}$ and the standard deviation of the error is also slightly higher .018 versus .011 for $\text{PCS}_{\text{Noise}}$. For SSL and all single-strategies, the strategy application-error ϵ is higher (up to .39 for TTB). $\text{PCS}_{\text{Transf}}$ predicts 77% of choices correctly, while $\text{PCS}_{\text{Noise}}$ and SSL perform worse with 72% and 70% (Table 3, lower part). PCS can explain 84% of all participants (Table 3, lower part) best: Most participants (57%) can be best explained by $\text{PCS}_{\text{Transf}}$ followed by 27% for $\text{PCS}_{\text{Noise}}$. A minority of 5% and 7% of participants can be best explained by SSL and the single strategy TTB. Mean posterior probabilities for $\text{PCS}_{\text{Transf}}$ and $\text{PCS}_{\text{Noise}}$ are high (.89 and .86). The overall likelihood for $\text{PCS}_{\text{Transf}}$ is high for $\text{PCS}_{\text{Transf}}$ with $\text{PCS}_{\text{Transf}}$ being e^{171} more likely than $\text{PCS}_{\text{Noise}}$ and e^{419} more likely than SSL.

6.3 Summary

As consequence of using tasks with increased difficulty, participants choose only only in 73% of all comparisons the better option. Increased difficulty in decision tasks also resulted in less model-consistent choice-behavior. The best model $\text{PCS}_{\text{Transf}}$ can explain 77% of all choices and therefore 13% to 14% less choices in comparison to the first and second study. Although PCS with a modified

Delta-Rule can explain most of the participants best, the PCS-implementation with a transformation function can handle the choices better than PCS with noisy learning. In the next section, we take a closer look at the behavior of models and how the free parameters affect predicted choice-probabilities that also provides a hint why $\text{PCS}_{\text{Transf}}$ is better than $\text{PCS}_{\text{Noise}}$ in the third study.

7. Overall model evaluation

7.1 Predicting choices

Predicted choice-probabilities of a typical participant from the first study show how $\text{PCS}_{\text{Transf}}$, $\text{PCS}_{\text{Noise}}$, and SSL behave (Figure 5, upper left panel). For $\text{PCS}_{\text{Transf}}$, there is low variance in choice-probabilites for correctly and incorrectly predicted choices: All correctly versus incorrectly predicted choices have the same probability around .90 versus .10. Those boundaries can be shifted by the sensitivity parameter γ: When γ decreases, boundaries also decrease as can be seen in comparison to another participant from the third study with a lower $\gamma = 1.345$ versus $\gamma = 2.195$ (Figure 6, upper left panel). When γ approximates zero, both boundaries approximate guessing probabilities at .5 (see Equation (7)). The reason that $\text{PCS}_{\text{Transf}}$ does not discriminate within (in-)correctly predicted choices is that activations of nodes in PCS only differ to a small extent between trials due to the tendency of the network to maximize differences in option-nodes for coherence in the network. These small difference do not have a significant effect when transformed into probabilities by an exponential choice rule (Equation (8)). $\text{PCS}_{\text{Noise}}$ behaves differently: Predicted choice-probabilities for the observed choices are scattered across the whole range from 0 to 1. Thus, the error implementation in $\text{PCS}_{\text{Noise}}$ gives a more differentiated prediction for choice-probablities. For both models, the learning rate is lower for a typical participant in the first study than in the second study (compare panel A in Figure 5 and Figure 6) in which participants tended to perform worse (Figure 3).

SSL behaves similar to $\text{PCS}_{\text{Noise}}$ with predicted choice-probabilities around .90 and .10. When both strategies TTB and WADD make the same correct or incorrect choice-predictions, the choice-probabilities are identical to the application-error of the strategies ϵ. When strategies make different predictions, initial preferences for strategies β as well as the learning rate w affect choice probabilities. In the example of a typical participant from study 1 (Figure 5, panel A), the participant has an initial higher preference for WADD (i.e., $\beta_{\text{WADD}} = 1 - \beta_{\text{TTB}} = .76$). Whenever TTB makes an incorrect prediction and WADD a correct prediction for this participant, choice probabilities are above .5 close to the initial preference (the exact size is further moderated by ϵ). Whenever WADD makes an incorrect prediction and TTB a correct prediction, choice-probabilities are below .5. Since this participant has a high learning rate (i.e., a low tendency to stick to the initial preferences for strategies), the participant tends to increase her preference for WADD as seen in an increasing curve (Figure 5, panel A). If the learning rate is low (i.e., the tendency to stick to a strategy high), probabilities stay constant (Figure 6, panel A).

Scatterplots for predicted choice probabilities for all studies between the three models are plotted jointly for the first two studies (Figure 5, panel B to D) for which results are very similar and separately for the third study (Figure 6, panel B to D). For the first two studies, there are 85% to 88% overlap of the same choices correctly predicted in each model comparison (i.e., perccentage of choices in the upper right quadrant in panel B: $\text{PCS}_{\text{Transf}}$ versus $\text{PCS}_{\text{Noise}}$, panel C: $\text{PCS}_{\text{Transf}}$ versus SSL, and panel D: $\text{PCS}_{\text{Noise}}$ versus SSL) . 7% to 8% of the same choices are incorrectly predicted by the models (lower left quadrant in each panel). Differences of unique correctly and incorrectly predicted choices (i.e., differences in percentages of choice in the upper left and lower right quadrant of each panel) range from 1% (comparison between $\text{PCS}_{\text{Transf}}$ and $\text{PCS}_{\text{Noise}}$ in panel B and $\text{PCS}_{\text{Noise}}$ and SSL in panel D) to 2% (comparison between $\text{PCS}_{\text{Transf}}$ and SSL in panel C). The pattern observed for a single participant as discussed can be also observed overall: For $\text{PCS}_{\text{Transf}}$ 99% of all predicted choice-probabilities are in the range of $.80$ to 1 for correct decision and .2 to 0 for incorrect predictions. $\text{PCS}_{\text{Noise}}$ and SSL scatter choice-predictions across the entire range (panel D) although for different reasons as discussed for the typical single participants. That is, for $\text{PCS}_{\text{Noise}}$ the variance of choice-probabilities can be observed within a single participant whereas the variance of the choice-probabilities over all participants result from four different lines of choice probabilities for each participant. For the third study (Figure 6) the pattern of results is more extreme: Differences in unique (in-)correctly predicted choices range from 2% (panel D) to 6% (panel C). Additionally, $\text{PCS}_{\text{Noise}}$ still predicts choices with a high probability (i.e., 15% in the range of .90 to 1) whereas $\text{PCS}_{\text{Transf}}$ and SSL almost never (3% of all choices) predicts extreme probabilities for choices (see most right column number in panel B to D). Put differently, $\text{PCS}_{\text{Noise}}$ makes more bold predictions whereas $\text{PCS}_{\text{Transf}}$ and SSL tend to decrease the number of

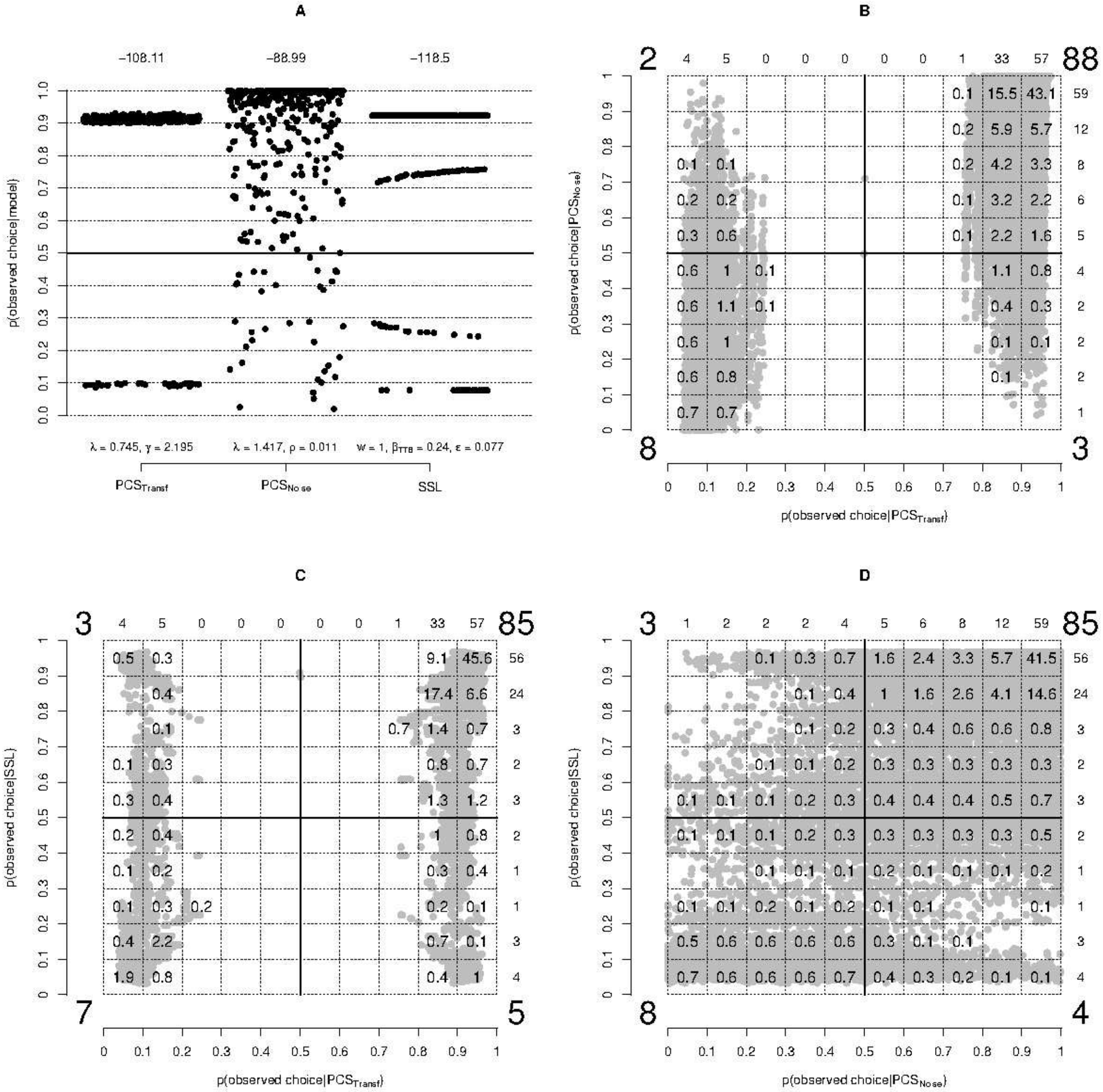

Figure 5. Choice-probabilities and sum of log-Likelihoods for observed choices as predicted by PCS_{Noise}, PCS_{Transf}, and SSL for a typical participant (panel A) and for all 175 participants$\times 363$ trials $=$ $63{,}525$ observed choices in Study 1 and 2 plotted between models (gray dots, panel B to D). In panel A, choice-probabilities for each of the three models are plotted from left to right for each model in the order of the decisions made by the participant. In panel B to D, numbers in the corners indicate the percentage of observed choices predicted to lie in each of the four quadrants as indicated by the thick black lines. Small numbers at the columns and rows of the plot indicate the sums of choice-probabilities for each model over the column- or row-cells of the other model. Small numbers in the cells indicate percentage of observed choices predicted to lie in cells of a range of .1 as indicated by gray dotted lines.

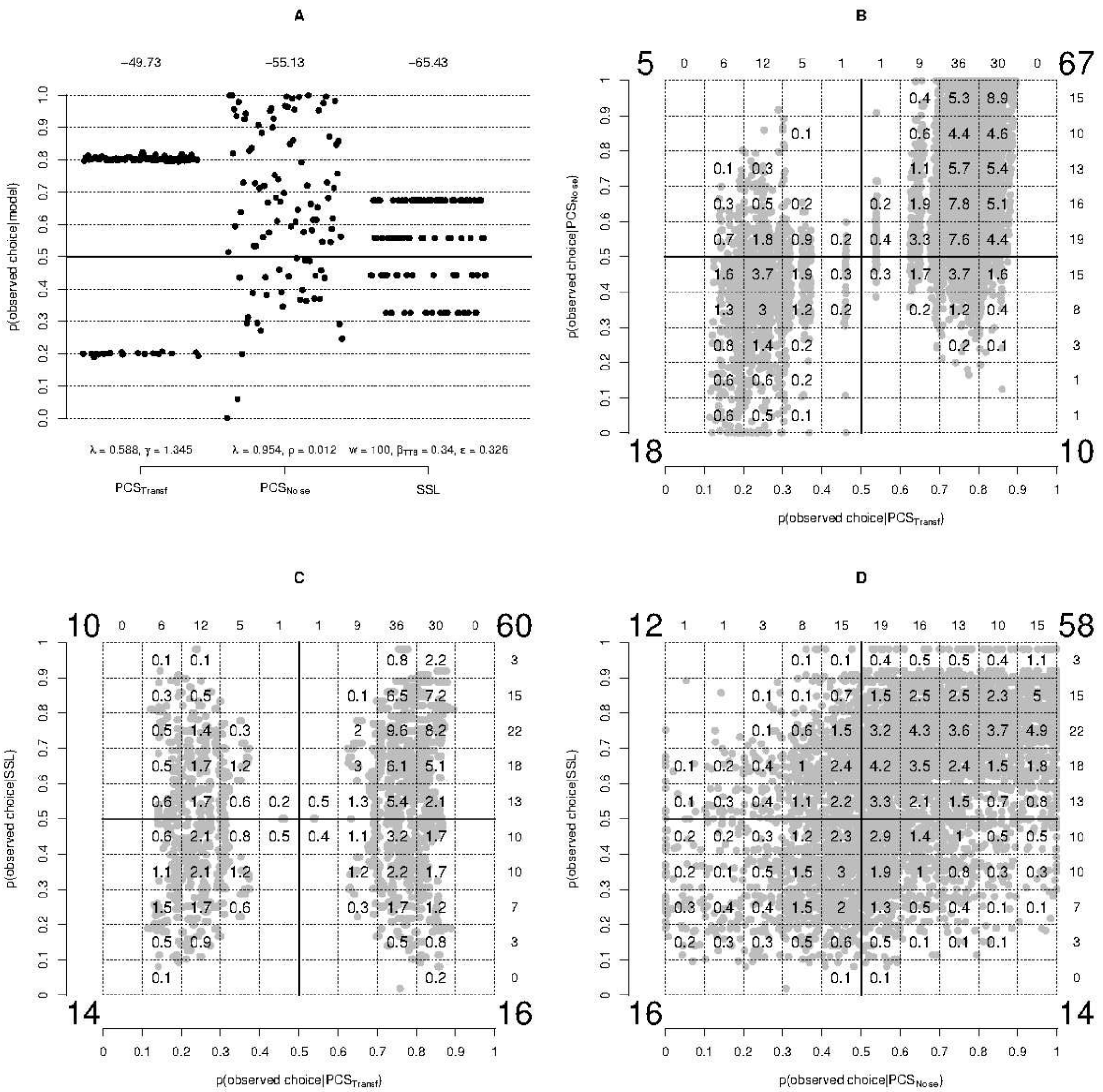

Figure 6. Choice-probabilities and sum of log-Likelihoods for observed choices as predicted by $\text{PCS}_{\text{Noise}}$, $\text{PCS}_{\text{Transf}}$, and SSL for a typical participant (panel A) and for all 60 participants $\times$ 100 trials $=$ $6{,}000$ observed choices in Study 3 plotted between models (panel B to D, gray dots). See the conceptually identicalFigure 5 for a detailed description of the four panels.

extreme predictions by lowering sensitivity γ to differences in activations in $\text{PCS}_{\text{Noise}}$ and by increasing strategy-application error ϵ in SSL.[4]

[4]Choice-probabilities for a typical participant plotted for the third study (Figure 6) is also illustrative: When removing the two choices for which $\text{PCS}_{\text{Noise}}$ has the lowest predicted choice-probabilities (i.e., 0.001 and 0.059), $\text{PCS}_{\text{Noise}}$ has the highest log-Likelihood with -45.39 versus -46.45 for $\text{PCS}_{\text{Transf}}$ and -63.79 for SSL. Thus, which model fits the data per participant best is sometimes determined by two extreme but incorrect choice-predictions.

7.2 Validating model parameters

In the final analysis, we evaluate the conceptual validity of the free parameters in $\mathrm{PCS_{Transf}}$, $\mathrm{PCS_{Noise}}$, and SSL by relating model-parameters to the overall performance of participants (i.e., relative number of correct decisions). For $\mathrm{PCS_{Transf}}$ we predict that an increasing learning rates λ and sensitivity for differences in evidence as modelled with the parameter γ is associated with an increasing performance. For $\mathrm{PCS_{Noise}}$ we expect the same for the learning rate λ and a negative relation between the standard deviation of the error ρ and performance (i.e., lower performance when learning is noisy). For SSL we predict that a higher error ϵ in the application of strategies leads to lower performance. To test the predictions, we ran four linear regressions (Table 4). Environmental factors (i.e., differences in validities and consequences) can explain 71% of the variance in participants' performance in the baseline model (first row). Results from the regressions with controls and model-parameters show support for all predictions: A higher learning rate in both PCS implementations lead to higher performance rates (second and third row, first column for parameters), a high sensitivity γ and a low standard deviation for error are significantly related to higher performance (second and third row, second column for parameters), and a lower error rate ϵ in SSL is related to a higher performance (fourth row, 3 column for parameters). A comparison between explained variance reveals (last column) that model parameters can explain an additional 11% to 14% of unique variance in participants' performance beyond controls.

Table 4. Regression-weights and explained variance from four linear regressions (table-rows) predicting participants' performance based on environmental characteristics of the task as controls (i.e., cue-validities, positive/negative reinforcement) in the base model (first line) and base model plus individual model parameters for $\mathrm{PCS_{Transf}}$ (second line), $\mathrm{PCS_{Noise}}$ (third line), and SSL (fourth line) for all three studies.

		Controls				Parameters			
	Int	Noncomp 1	Comp 2	Comp 3	Pos	(1:λ, 2:λ, 3:w)	(1:γ, 2:ρ, 3:β_{TTB})	(3:ϵ)	R^2
Baseline	0.887***	0.01	0.024*	-0.151***	-0.013*	—	—	—	.711
$\mathrm{PCS_{Transf}}$	0.67***	0.006	-0.006	-0.062***	-0.008	1: 0.027***	1: 0.096***	—	.846
$\mathrm{PCS_{Noise}}$	0.877***	0.011	0.013	-0.096***	-0.006	2: 0.043***	2: -4.646***	—	.838
SSL	0.92***	0.014	0.024*	-0.076***	-0.009	3: <.001	3: 0.013	3: -0.528***	.819

Note. Significance: $p^* < .05$, $p^{**} < .01$, $p^{***} < .001$

7.3 Comparing both PCS-implementations

In the article, we introduced two PCS-implementations of learning ($\text{PCS}_{\text{Transf}}$ and $\text{PCS}_{\text{Noise}}$) that differ in how predictions are derived (i.e., sensitivities to differences in activation or noisy learning) but assume the same learning-algorithm (Delta-rule). Both PCS models make different predictions for decision times. The number of iterations on $\text{PCS}_{\text{Transf}}$ and the mean number (i.e., averaged over 1000 simulated participants) of iterations in $\text{PCS}_{\text{Noise}}$ are a thereby used as a proxy for decision time: The higher the number of iterations I, the slower the expected decision time. Second, $\text{PCS}_{\text{Noise}}$ differs from all other models in its tendency to scatter probabilities for predicted choices across the entire range of probabilities (Figure 5). We show in the following that this behavior of $\text{PCS}_{\text{Noise}}$ is sensitive to characteristics of the participant in the decision-environment and thereby accounts for meaningful variance in the data.

7.3.1 Predicting decision times

On average over all participants and tasks in all three studies, participants need 1.93 seconds ($SD = 2.04$; Median $= 1.30$) to make a decision. Individual correlations between decision times and predictions for each participant are higher for $\text{PCS}_{\text{Noise}}$ with 50% of all correlations between $r = .23$ and $r = .38$ and a mean correlation of $r = .30$ for $\text{PCS}_{\text{Noise}}$ versus $r = .24$ for $\text{PCS}_{\text{Transf}}$ (Figure 7). A comparison between two multilevel-models with either predictions of $\text{PCS}_{\text{Transf}}$ or $\text{PCS}_{\text{Noise}}$ included as predictors for decision time (with a random intercept and slope conditioned on participants) shows overwhelming evidence for $\text{PCS}_{\text{Noise}}$ for explaining decision times with a Bayes Factor of e^{664} in comparison to $\text{PCS}_{\text{Transf}}$. Differences in predictions of decision times between PCS-models is explained in more detail in Appendix D.

7.3.2 Validity of variance in predicted choice-probabilities for PCS with noisy learning

$\text{PCS}_{\text{Noise}}$ shows a high variance for predicted choice-probabilities (e.g., upper left panel Figure 5). In the following, we test whether predicted-choice probabilities relate to the data in a meaningful way. Choice-probabilities are derived from $\text{PCS}_{\text{Noise}}$ by simulating 1000 participants that differ only in random noise added to the net-weights representing subjective validities for experts. For example, a

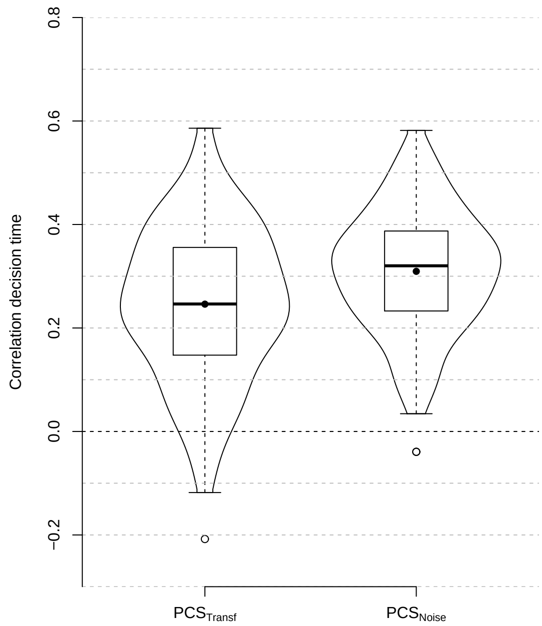

Figure 7. Distributions of individual correlations between observed decision-times and predicted decision times for $\text{PCS}_{\text{Transf}}$ and $\text{PCS}_{\text{Noise}}$.

probability of .55 for a choice for option A results from 550 of the 1000 simulated participants choosing option A and 450 participants choosing option B. Predicted choice probabilities from $\text{PCS}_{\text{Noise}}$ are influenced by the difficulty of the trial: When the evidence for options is similar, small differences in the representation of net-weights due to noise result in choice-probabilities closer to .5. For easy trials (clear evidence for one option), noise does not alter the decision of the simulated participants considerably. Thus, the difficulty of the trials for real participants should be correlated with the predicted choice probabilities from the model: Increasing predicted choice-probabilities for $\text{PCS}_{\text{Noise}}$ should correspond to increasing probabilities for participants choosing the better option. We ran a multilevel logistic-regression (allowing for a random intercept and slope for each participant) including participants from all studies: We predicted the criterion of an observed correct choice (i.e., coded as 1 = correct and 0 = incorrect) and the maximal choice-probability for options in each trial t derived from $\text{PCS}_{\text{Noise}}$ as predictor (i.e., $max[p(opt_{c,I,t}); p(opt_{nc,I,t})]$). We find the predicted relation: When predicted choice-probabilities increase by one percent (e.g., .71 to .72), the odds of a participant solving a trial increase by a factor of 1.06 ($p < .001$). The high variance in choice-probabilities for $\text{PCS}_{\text{Noise}}$ corresponds to the variance in probabilities of making a correct decision as predicted which further speaks for the validity of PCS_{Noise}.

7.4 Overall summary

In three studies with 235 participants from two independent participant-pools making 69,525 choices in four different environments with different types of consequences of decisions (rewards versus punishment), the overall evidence for learning in PCS versus the adaptive toolbox is overwhelming (Table 3, lower part). The Bayes-factor for the best fitting PCS-model (i.e., PCS_{Noise}) and the best fitting model from the adaptive toolbox (i.e., SSL) speaks with e^{2044} in favor of PCS. Most participants (94%) can be best explained by PCS. Additionally, PCS predicts 89% whereas SSL predicts 86% of all choices correctly; this means, that PCS makes $.03 \times 69{,}525 \simeq 2085$ more correct choices.[5] Analysis of participants' decision times show that PCS with noisy learning can explain decision times better than PCS with a transformation function speaking for PCS_{Noise} overall as a more accurate model of the decision-process.

8. Theoretical implications

Results show that the assumption of powerful automatic cognitive processes that take many pieces of information into account can explain human learning of cue-criterion relations and integration better than a prototypical single (non-)compensatory strategy or a mix of both prototypes as specified in SSL. People are able to learn cue-criterion relations and the interrelation between cues and also track changes in those relations which are all aspects of the network-model (i.e., the entire available information is represented as cue-nodes and validity-weights and their interrelation is modeled in the interconnections between cue-nodes and option-nodes). Our results do not support the claim from the adaptive toolbox that people simplify the task by strategically reducing the amount of information as formulated in the typical non-compensatory strategy Take-the-best or by ignoring the interrelation of information as specified in WADD. Low decision times (i.e., between one and two seconds) and mostly incorrect rankings of cues according to validity further speak for automatic processes that resemble characteristics of intuition (Betsch & Glöckner, 2010).

Dual-process theories of thinking (e.g., Hogarth, 2001, 2005; Kahneman, 2011; Kahneman & Klein, 2009) assume a distinction between an automatic, fast, unintentional, and effortless mode (tacit system

[5] In Appendix E, we show that conclusions are stable in a cross-validation).

or system 1) and a deliberate, analytical, and limited mode of thinking that requires effort (deliberate system or system 2). It has been argued that heuristics are "shortcuts to deliberation" (Betsch & Glöckner, 2010, p. 286) working within system 2 whereas unintentional processes as modeled in PCS are parts of system 1. Learning within these two systems may differ. Learning from (e.g.) others (Hertwig, Hoffrage, & The ABC Research Group, 2013) to apply a heuristic that is smart for a domain (e.g., learning of cues for trustworthiness: Gambetta & Hamill, 2005) may demand attention and effort from system 2 until heuristics are internalized and can easily be retrieved as automatized production-rules (Marewski & Mehlhorn, 2011) whereas learning from experience in system 1 "depends heavily on automatic processes [as modeled in PCS] and is therefore much easier" (Hogarth, 2001, p. 189).

Finally, heuristics may also influence learning in system 1 indirectly by restructuring the decision-situation through searching, adding, and generating new information in case the initial automatic response does not lead to a coherent interpretation of the decision-situation (Betsch & Glöckner, 2010; Glöckner & Betsch, 2008). Overall, the studies show that learning by experience can be best modeled by a parallel-constraint satisfaction network model that captures automatic processes of learning. How automatic learning process can be influenced by heuristics (e.g., by restructuring the network) and how this interplay between the two systems can be modeled in a formalized meta-theory is an open question for further integrative research.

References

Adelman, L. (1981). The influence of formal, substantive, and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. *Organizational Behavior and Human Performance, 27*, 423–442.

Balzer, W. K., Sulsky, L. M., Hammer, L. B., & Sumner, K. E. (1992). Task information, cognitive information, or functional validity information: Which components of cognitive feedback affect performance? *Organizational Behavior and Human Decision Processes, 53*, 35–54.

Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review, 3*, 439-449.

Bechtel, W., & Abrahamsen, A. (2002). *Connectionism and the mind*. Malden: Blackwell Publishing.

Betsch, T., & Glöckner, A. (2010). Intuition in judgment and decision making: Extensive thinking without effort. *Psychological Inquiry, 21*, 279–294.

Brehmer, B. (1969). Cognitive dependence on additive and configural cue-criterion relations. *The American Journal of Psychology, 82*, 490–503.

Brehmer, B. (1980). In one word: Not from experience. *Acta psychologica, 45*, 223–241.

Brownstein, A. L., Read, S. J., & Simon, D. (2004). Bias at the racetrack: Effects of individual expertise and task importance on predecision reevaluation of alternatives. *Personality and Social Psychology Bulletin, 30*, 891–904.

Bröder, A. (2000a). "Take The Best - Ignore The Rest": Wann entscheiden Menschen begrenzt rational? Lengerich: Pabst.

Bröder, A. (2000b). Assessing the empirical validity of the "Take-The-Best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26* (5), 1332-1346.

Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 611-625.

Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 904-918.

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217.

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological Review, 118*, 193–218.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review, 100*, 432–459.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing, 16*, 1190–1208.

Chasseigne, G., Mullet, E., & Stewart, T. R. (1997). Aging and multiple cue probability learning: The case of inverse relationships. *Acta Psychologica, 97*, 235–252.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*, 181–204.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 97-118). New York: Oxford University Press.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & Van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311*, 1005–1007.

Dijksterhuis, A., & Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological science, 1*, 95–109.

Doherty, M. E., & Brehmer, B. (1997). The paramorphic representation of clinical judgment: A thirty-year retrospective. In W. M. Goldstein & R. M. Hogarth (Eds.), (pp. 537-551). New York: Cambridge University Press.

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological plausibility of the theory of probabilistic mental models and the fast and frugal heuristics. *Psychological Review, 115*, 199–211.

Dougherty, M. R., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106*, 180–209.

Eddelbuettel, D., & Francois, R. (2011). Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software, 40*, 1-18.

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases.

Psychological Review, 107, 659–676.

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 186–203.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*, 247–279.

Gambetta, D., & Hamill, H. (2005). *Streetwise: How taxi drivers establish their customers' trustworthiness*. New York: Russell Sage Foundation.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science, 1*(1), 107-143.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650-669.

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Fast and frugal heuristics: Theory, tests, and applications*. Oxford: University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Gigerenzer, G., Hoffrage, U., & Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: Reply to Dougherty, Franco-Watkins, and Thomas (2008). *Psychological Review, 115*, 230–239.

Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Glöckner, A. (2009). Investigating intuitive and deliberate processes statistically: The multiple-measure maximum likelihood strategy classification method. *Judgment and Decision Making, 4*, 186–199.

Glöckner, A. (2010). Multiple measure strategy classification: Outcomes, decision times and confidence. In A. Glöckner & C. L. M. Witteman (Eds.), *Tracing intuition: Recent methods in measuring intuitive and deliberate processes in decision making* (pp. 83-105). London: Psychology Press & Routledge.

Glöckner, A., & Betsch, T. (2008). Modeling option and strategy choices with connectionist networks: Towards an integrative model of automatic and deliberate decision making. *Judgment and Decision Making, 3*, 215–228.

Glöckner, A., & Betsch, T. (2012). Decisions beyond boundaries: When more information is processed faster than less. *Acta pPsychologica, 139*, 532–542.

Glöckner, A., Betsch, T., & Schindler, N. (2010). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making, 23*, 439–462.

Glöckner, A., & Bröder, A. (2011). Processing of recognition information and additional cues: A model-based analysis of choice, confidence, and response time. *Judgment and Decision Making, 6*, 23–42.

Glöckner, A. & Bröder, A. (2014). Cognitive integration of recognition information and additional cues in memory-based decisions. *Judgment and Decision Making, 9*, 35–50.

Glöckner, A., Heinen, T., Johnson, J. G., & Raab, M. (2012). Network approaches for expert decisions in sports. *Human Movement Science, 31*, 318–333.

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction theory. *Cognition, 133*, 641–666.

Glöckner, A., & Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgment and decision making. *Thinking & Reasoning, 16*, 1–25.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*, 227–247.

Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer & V.Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63* (pp. 79–93). Göttingen: Ges. für Wiss. Datenverarbeitung.

Hammond, K. R., & Stewart, T. R. (2001). *The essential Brunswik: Beginnings, explications, applications*. Oxford University Press.

Hammond, K. R., Summers, D. A., & Deane, D. H. (1973). Negative effects of outcome-feedback in multiple-cue probability learning. *Organizational Behavior and Human Performance, 9*, 30–34.

Hertwig, Hoffrage, & The ABC Research Group. (2013). *Heuristics in a social world.* New York: Oxford University Press.

Hirst, M. K., & Luckett, P. F. (1992). The relative effectiveness of different types of feedback in performance evaluation. *Behavioral Research in Accounting, 4*, 1–22.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin, 57*, 116–131.

Hoffman, P. J., Earle, T. C., & Slovic, P. (1981). Multidimensional functional learning (MFL) and some new conceptions of feedback. *Organizational Behavior and Human Performance, 27*, 75–102.

Hogarth, R. M. (2001). *Educating intuition*. Chicago: University Press.

Hogarth, R. M. (2005). Deciding analytically or trusting your intuition? The advantages and disadvantages of analytic and intuitive thought. In Betsch, T., & Haberstroh, S. (Eds.), *The routines of decision making* (pp. 67–82). New Jersey: Taylor & Francis.

Hogarth, R. M. (2010). Intuition: A challenge for psychological research on decision making. *Psychological Inquiry, 21*, 338–353.

Hogarth, R. M., & Karelaia, N. (2005). Simple models for multiattribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Science, 51*, 1860–1872.

Hogarth, R. M., & Karelaia, N. (2006a). Regions of rationality: Maps for bounded agents. *Decision Analysis, 3*, 124–144.

Hogarth, R. M., & Karelaia, N. (2006b). "Take-the-Best" and other simple strategies: Why and when they work "well" with binary cues. *Theory and Decision, 61*, 205–249.

Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review, 114*, 733–758.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*, 3–31.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295–355.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences, USA, 79*, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded responses have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, USA, 81*, 3088–3092.

Jekel, M., Fiedler, S., & Glöckner, A. (2011). Diagnostic task selection for strategy classification in judgment and decision making: Theory, validation, and implementation in R. *Judgment and Decision Making, 6*, 782–799.

Jekel, M., Glöckner, A., Fiedler, S., & Bröder, A. (2012). The rationality of different kinds of intuitive decision processes. *Synthese, 189*, 147–160.

Johnson, T. R., Zhang, J., & Wang, H. (1997). A hybrid learning model of abductive reasoning. In R.

Sun & F. Alexandre (Eds.), *Connectionist-symbolic integration: From unified to hybrid approaches* (pp. 91–112). Mahwah, NJ: Lawrence Erlbaum.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Penguin.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*, 515–526.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237–251.

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin, 134*, 404–426.

Katsikopoulos, K. V., Schooler, L. J., & Hertwig, R. (2010). The robust beauty of ordinary information. *Psychological Review, 117*, 1259–1266.

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.

Marewski, J. N., & Mehlhorn, K. (2011). Using the ACT-R architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making, 6*, 439–519.

Marewski, J. N., & Olsson, H. (2009). Beyond the null ritual. *Journal of Psychology, 217*, 49–60.

Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 119-140). New York: Oxford University Press.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences, 14*, 348–356.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basics findings. *Psychological Review, 88*, 375-407.

McClelland, J. L., Rumelhart, D. E., & The PDP Research Group. (1986). *Parallel distributed processing. Explorations in the microstructure of cognition, (Volume 2: Psychologocial and biological models)*. Cambridge: MIT Press.

Monroe, B. M., & Read, S. J. (2008). A general connectionist model of attitude structure and change: the ACS (Attitudes as Constraint Satisfaction) model. *Psychological Review, 115*, 733–759.

Muchinsky, P. M., & Dudycha, A. L. (1975). Human inference behavior in abstract and meaningful environments. *Organizational Behavior and Human Performance, 13*, 377–391.

Myung, I. J., Pitt, M. A., & Kim, W. (2005). Model evaluation, testing and selection. In K. Lamberts & R. L. Goldstone (Eds.), *Handbook of Cognition* (pp. 422–436). London: Sage Publications.

Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences, 9*, 11–15.

Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors influencing "one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 53–65.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*, 534–552.

Pinheiro, J.C., & Bates, D. M. (2000). *Mixed-Effects models in S and S-PLUS*. Heidelberg: Springer.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review, 117*, 864–901.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing.* Vienna: Austria.

Read, S. J., Monroe, B. M., Brownstein, A. L., Yang, Y., Chopra, G., & Miller, L. C. (2010). A neural network model of the structure and dynamics of human personality. *Psychological review, 117*, 61–92.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1*, 26–53.

Reilly, B. A., & Doherty, M. E. (1992). The assessment of self-insight in judgment policies. *Organizational Behavior and Human Decision Processes, 53*, 285–309.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 65-99). New York: Appleton-Century-Crofts.

Rieskamp, J. (2006). Perspectives of probabilistic inferences: Reinforcement learning and an adaptive network compared. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 1355–1370.

Rieskamp, J. (2008). The importance of learning when making inferences. *Judgment and Decision Making, 3*(3), 261–277.

Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of*

Experimental Psychology: General, 135, 207–236.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533–536.

Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. *Parallel distributed processing: Explorations in the microstructure of cognition (Volume 1: Foundations)*. Cambridge: MIT Press.

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

Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review, 103*, 219–240.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*, 1–19.

Simon, D., Snow, C. J., & Read, S. J. (2004). The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology, 86*, 814–837.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology, 70*, 893–912.

Söllner, A., Bröder, A., Glöckner, A., & Betsch, T. (2014). Single-process versus multiple-strategy models of decision making: Evidence from an information intrusion paradigm. *Acta Psychologica, 146*, 84–96.

Thagard, P. R. (1988). *Computational philosophy of science*. Cambridge: MIT Press.

Thagard, P. R. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12*, 435–502.

Thagard, P. (2003). Why wasn't OJ convicted? Emotional coherence in legal inference. *Cognition & Emotion, 17*, 361–383.

Thagard, P. R., & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal-driven learning* (pp. 439-454). Cambridge: MIT Press.

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115*, 155–185.

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

Trueblood, J., Brown, S.D., & Heathcote, A. (in press). The multi-attribute linear ballistic accumulator model of context effects in multi-alternative choice. *Psychological Review*.

Vanhoomissen, T., & Van Overwalle, F. (2010). Me or not me as source of ingroup favoritism and outgroup derogation: a connectionist perspective. *Social Cognition, 28*, 84–109.

Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74*, 312–328.

Van Overwalle, F. (2007). Social connectionism: A reader and handbook for simulations. New York: Psychology Press.

Van Overwalle, F. V., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review, 9*, 231–274.

Zacks, R. T., & Hasher, L. (2002), Frequency processing: A twenty-five year perspective. In P. Sedlmeier & T. Betsch (Eds.), *Etc.: Frequency processing and cognition* (pp. 21–36). New York: Oxford University Press.

APPENDICES

A. Models from the adaptive toolbox

A.1 Updating cue-validities

The validity of cue l at trial $t = \{1 \ldots T\}$ is defined as a relative frequency of past trials in which the cue signaled the better option divided by the trials in which the cue discriminated between choice-options:

$$val_{cue_l,t} = \frac{f(\text{correct}_{t-1})}{f(\text{correct}_{t-1}) + f(\text{incorrect}_{t-1})}. \tag{A.1}$$

For example, a cue that pointed towards the better option in 80 past trials and the worse option in 20 trials has a validity of $val_{cue_1,t} = \frac{80}{80+20} = .80$ at the present trial t.

A.2 Single strategies

In the model-comparison, we consider $P = 3$ prominent single strategies s_p. The first model is the non-compensatory strategy Take-the-best (Gigerenzer & Goldstein, 1996; Todd, Gigerenzer, & The ABC Research Group, 2011). A participant who uses Take-the-best (TTB) inspects cues in the order of validity and decides in line with the first cue that discriminates between options. Another prominent compensatory strategy is Weighted Additive (Gigerenzer & Goldstein, 1996; Todd, Gigerenzer, & The ABC Research Group, 2011). A participant who uses Weighted Additive (WADD) sums the weighted cue pattern for each option and picks the option with the highest evidence $EV_{opt_{k,t}}$ according to:

$$EV_{opt_{k,t}} = \sum_{l=1}^{L} val_{cue_l,t} \times cue_{l,t} \tag{A.2}$$

Finally, another compensatory strategy—the rational model (RAT)—is the integration of the probabilistic information for each comparison in line with (naïve) Bayes (Lee & Cummins, 2004). A participant who uses RAT sums the cue pattern weighted by log-odds for each option and picks the option with the highest evidence $EV_{opt_{k,t}}$ according to:

$$EV_{opt_{k,t}} = \sum_{l=1}^{L} ln(\frac{val_{cue_l,t}}{1 - val_{cue_l,t}}) \times cue_{l,t} \tag{A.3}$$

In line with past literature, we assume that participants may not always correctly apply a strategy but only with a trembling-hand error ϵ that is estimated from the choices of a participant. That is, the maximum likelihood of all correctly predicted choices n_{co}, incorrectly predicted choices n_{inco}, and predicted guessing n_{guess} by strategy s_p can be calculated by finding the individual parameters ϵ in the interval $[.001, .499]$ that maximizes:

$$ln(L_{s_p}) = ln[(1-\epsilon)^{n_{co}} \times \epsilon^{n_{inco}} \times .5^{n_{guess}}]. \tag{A.4}$$

A.3 SSL: Selecting strategies from the toolbox

According to the Strategy Selection Learning theory[6] (Rieskamp, 2006, 2008; Rieskamp and Otto, 2006), people can choose a strategy s_p from a set of P strategies when solving a probabilistic decision task. In line with prior work, we set $s_p = \{\text{TTB}, \text{WADD}\}$, that is, participants can either apply Take-the best (TTB) or Weihghted Additive (WADD). When participants start with the first decision task, they may differ in their initial preference for applying TTB or WADD. The parameter β_{TTB} ranging from 0 to 1 and WADD with $\beta_{\text{WADD}} = 1 - \beta_{\text{TTB}}$ accounts for this individual differences. Participants may also differ in the extent to which they are influenced by new experiences during the experiment versus how strong they stick to their initial preferences. This is accounted for by the parameter w, with $1 \leq w \leq 100$. A high w results in a slow learning rate. Following, $q_{t=1,s_p}$, the expectancy of strategy s_p for trial $t = 1$, can be calculated (cf. Rieskamp, 2006, Equation 2, p. 1356):

$$q_{t=1}(s_p) = r_{correct} \times w \times \beta_{s_p}. \tag{A.5}$$

The constant $r_{correct}$ is the potential payoff received in the first round and is used for scaling purposes only given payoffs change between tasks (Rieskamp, 2006, p. 1356); we therefore set $r_{correct} = 1$ in all studies.[7] Expectancies for the success of a strategy are transformed into probabilities for applying a strategy (cf. Rieskamp, 2006, Equation 1, p. 1356):

$$p_t(s_p) = \frac{q_t(s_p)}{\sum_k^{K=2} q_t(s_p)}. \tag{A.6}$$

That is, the probability of applying strategy s_p is the expectancy of this strategy normalized by the

[6] The following description of SSL is partially copied from Glöckner, Hilbig, and Jekel (2014).
[7] Conclusions do not depend on the exact value.

sum of expectancies of all strategies in order to receive probabilities that range between 0 and 1. The expectancy of strategy s_p is updated in the next and all following T decision tasks by (cf. Rieskamp, 2006, Equation 3, p. 1356):

$$q_{t>1}(s_p) = q_{t-1}(s_p) + I_{t-1}(s_p) \times r_{t-1}(s_p). \tag{A.7}$$

That is, the expectancy of strategy s_p being successful in the current trial $t > 1$ is the sum of the expectancy of the strategy for the previous decision trial $t - 1$ and the payoff received for the previous decision trial r_{t-1} multiplied with an indicator $I_{t-1}(s_p)$. The indicator $I_{t-1}(s_p)$ is 1 when the participant's choice was in accord with the prediction of the strategy in the previous trial $t - 1$. If the participant did not choose the option predicted by the strategy, the indicator is coded as $I_{t-1}(s_p) = 0$ and, thus, the expectancy for strategy s_p does not change. If both strategies make the same prediction and the participant decided in line with the strategies, $I_{t-1}(s_p)$ equals the probability predicted for the selection of the strategy, that is, $I_{t-1}(s_p) = p_t(s_p)$. The expectancy of a strategy can only take positive values. In case the expectancy of a strategy is below 0, the expectancy is set to a value of .0001 (Rieskamp, 2006, p. 1356). If strategy s_p predicts a choice for option $k = 1$ for trial t, the probability for a choice for option 1 for strategy s_p is $p(opt_{1,t}|s_p) = 1$; following, the probability for the alternative option 2 is $p(opt_{2,t}|s_p) = 0$. Allowing for an error ϵ in the application of strategies, the probability for a decision for option $k = 1$ given strategy s_p and error ϵ, with $0 < \epsilon < .5$, is (cf. Rieskamp, 2006, Equation 4, p. 1357):

$$p(opt_{1,t}|s_p, \epsilon) = (1 - \epsilon) \times p_t(opt_{1,t}|s_p) + \epsilon \times p_t(opt_{2,t}|s_p). \tag{A.8}$$

Finally, the probability of a choice for option k independent of the strategy s_p is the product of the probability for the application of strategy s_p and the probability of a choice for option k given strategy s_p and ϵ summed over all P strategies (cf. Rieskamp, 2006, Equation 5, p. 1357):

$$p(opt_{k,t}) = \sum_{p=1}^{P} p_t(s_p) \times p(opt_{k,t}|s_p, \epsilon). \tag{A.9}$$

To receive the sum of the log-likelihoods for the choices observed for a participant, two matrices, $\mathbf{P}_{T,2}$ and $\mathbf{I}_{2,T}$, are multiplied. In the first column of matrix $\mathbf{P}_{T,2}$, logarithmic probabilities for a choice for option $k = 1$ are included; in the second column, all logarithmic probabilities for a choice for option

$k = 2$ are included. Matrix $\mathbf{I}_{2,T}$ indicates the choices of a participant: if a participant chooses option 1 in trial t, $c_t = 1$; if she chooses option 2, $c_t = 0$. Matrix $\mathbf{R}_{T,T}$ results from a matrix multiplication of matrix $\mathbf{P}_{T,2}$ and matrix $\mathbf{I}_{2,T}$:

$$\mathbf{R}_{T,T} = \mathbf{P}_{T,2} \times \mathbf{I}_{2,T} = \begin{bmatrix} ln(p(opt_{1,1})) & ln(1 - p(opt_{1,1})) \\ ln(p(opt_{1,2})) & ln(1 - p(opt_{1,2})) \\ \vdots & \vdots \\ ln(p(opt_{1,T})) & ln(1 - p(opt_{1,T})) \end{bmatrix} \times \begin{bmatrix} c_1 & c_2 & \dots & c_T \\ 1 - c_1 & 1 - c_2 & \dots & 1 - c_T \end{bmatrix}. \tag{A.10}$$

The maximum of the log-likelihood function ln(L_{SSL}) of the choices of a participant can be calculated by finding the individual parameters of w in the interval $[1, 100]$ (Rieskamp, 2006, p. 1362), β_i in the interval $[.001, .999]$, and ϵ in the interval $[.001, .5]$ that maximize the sum of the log-likelihoods in the diagonal of the matrix $\mathbf{R}_{T,T}$:

$$ln(L_{\text{SSL}}) = \sum_{t=1}^{T} \mathbf{R}_{t,t}. \tag{A.11}$$

B. Model comparison: Calculating Bayes-factors and posterior probabilities

To account for model flexibility, the Bayesian Information Criterion (BIC, Schwarz, 1978) is calculated for model $m_j = \{\text{PCS}_{\text{Transf}}, \text{PCS}_{\text{Noise}}, \text{SSL}, \text{TTB}, \text{WADD}, \text{RAT}\}$ from the log-likelihoods as defined in Equation (8) for PCS, Equation (A.11) for SSL, and Equation (A.4) for all single strategies:

$$\text{BIC}_{m_j} = -2 \times ln(L_{m_j}) + ln(T) \times p_{m_j}. \tag{B.1}$$

That is, the log-likelihood increases by the sum of the product of the logarithmic number of decision trials T used to find the optimal model parameters p_{m_j} with $p_{\text{PCS}_{\text{Noise}}} = 2$ (i.e., learning rate λ, standard deviation of the noise ρ), $p_{\text{PCS}_{\text{Transf}}} = 2$ (i.e., learning rate λ, exponent of the transformation function γ), and $p_{\text{SSL}} = 3$ (i.e., initial preference β, learning rate w, and application error ϵ), and $p_{\text{TTB}} = p_{\text{WADD}} = p_{\text{RAT}} = 1$ (i.e., application error ϵ). From the BIC scores, the Bayes-factor $\text{BF}_{m_{j,x}}$ resulting from a comparison between model j and $x \neq j$ can be calculated (Wagenmaker, 2007, p. 796, Equation 9):

$$\text{BF}_{m_{j,x}} = e^{[-.5\times(\text{BIC}_{m_j}-\text{BIC}_{m_x})]}. \tag{B.2}$$

Finally, the posterior probability for model $j = k$, that is, the probability of model k as the data generating mechanism under consideration of the observed choices D and under the assumption of equal prior probabilities for all models, can be calculated from the BIC_{m_j} values according to (cf. Wagenmakers, 2007, Equation 11, p. 797):

$$\Pr(m_j|D) = \frac{e^{[-\frac{1}{2}\times\text{BIC}_{m_k}]}}{\sum_{j=1}^{J} e^{[-\frac{1}{2}\times\text{BIC}_{m_j}]}}. \tag{B.3}$$

C. Implementation of learning in PCS in the software-package R

C.1 PCS

In the default exemplary cue-pattern t specified in the arguments for the function “PCS()” from line 2 to 13, a cue with a net-weight of $w_{cue_1,t} = .4$ speaks for option 1 (thus, $w_{cue_1-opt_1,t} = .01$ and $w_{cue_1-opt_2,t} = -.01$) and another cue speaks with a net-weight of $w_{cue_2,t} = .3$ for option 2 (thus, $w_{cue_2-opt_1,t} = -.01$ and $w_{cue_2-opt_2,t} = .01$).

```
1  PCS = function(
2    activ = c(1,0,0,0,0),   # node-activations at i=1: a_source a_cue1,1,t a_cue2,1,t a_opt1,1,t a_opt2,1,t
3    weightsNet = rbind(
4      c(0,.4,.3,0,0),       # w_source-source,t w_cue1,t w_cue2,t w_cue1-opt1,t w_cue1-opt2,t
5      c(.4,0,0,.01,-.01),   # w_cue1,t w_cue1-cue1,t w_cue1-cue2,t w_cue1-opt1,t w_cue1-opt2,t
6      c(.3,0,0,-.01,.01),   # w_cue2,t w_cue2-cue1,t w_cue2-cue2,t w_cue2-opt1,t w_cue2-opt2,t
7      c(0,.01,-.01,0,-.2),  # w_opt1-source,t w_opt1-cue1,t w_opt1-cue2,t w_opt1-opt1,t w_opt1-opt2,t
8      c(0,-.01,.01,-.2,0)), # w_opt2-source,t w_opt2-cue1,t w_opt2-cue2,t w_opt2-opt1,t w_opt2-opt2,t
9    flo = -1,
10   ceil = 1,
11   decay = .1,
12   stability = 10^6,
13   maxiter = 1000
14 ){
15   # define variables
16   ener = rep(0, maxiter)
17   iter = 1
18   # activation matrix at i = 1
19   matWeightedAct = activ * weightsNet
```

```
  while (iter != maxiter){
    # input for each node in the net; Equation (1)
    input = colSums(matWeightedAct)
    # activation for each node for i = 1 + iter; Equation (2)
    activ =  (1-decay) * activ + input *
      ifelse(round(input, 6) < 0, activ - flo, ceil - activ)
    # set activation of source node to 1
    activ[1] = 1
    # activation matrix at i = 1 + iter
    matWeightedAct = activ * weightsNet
    # energy at i = 1 + iter; Equation (3)
    ener[iter] = -sum(t(matWeightedAct) * activ)
    # evaluate stopping criterion
    if (iter > 10){
      if(sum(floor(stability * (ener[(iter  - 10) : (iter-1)])) -
               floor(stability * (ener[iter]))  == 0) == 10){
        iter = iter + 1
        break()
      }}
    iter = iter + 1
  }
  # prompt output of function
  return(c(iter, ener[iter-1], activ))
}
```

Note. Comments are in red and R-commands are in bold. The function is also implemented in Rcpp (Eddelbuettel & Francois, 2011) for faster execution and can be retrieved from the first author.

C.2 Learning in PCS

In the default exemplary cue-pattern t specified in the arguments for the function “deltaRule()” from line 2 to 7, the observed network-activation at the final iteration I is $a_{opt_1,I,t} = .50$ and $a_{opt_1,I,t} = -.50$ for option 1 and 2. Thus, the network proposes a decision for option 1. The decision is wrong (i.e., the network receives feedback that option 2 is the better option) as coded in a desired activation of $d_{a_{opt_1,I,t}} = -.6$ and $d_{a_{opt_2,I,t}} = .6$ for option 1 and 2. Given a default learning-rate of $\lambda = .2$, the function updates the network-weights for the cues by decreasing net-weight $w_{cue_1,t}$ and increasing net-weight $w_{cue_2,t}$.

```
1 deltaRule = function(
2   activ = c(.50,-.50),                       # a_{opt_1,I,t}  a_{opt_2,I,t}
3   desiredActiv = c(-.6, .6),                 # d_{a_{opt_1,I,t}}  d_{a_{opt_2,I,t}}
4   weightsCuesOptions = cbind(c(.4,.01,-.01), # w_{cue_1,t}  w_{cue_1-opt_1,t}  w_{cue_1-opt_2,t}
```

```
                             c(.3,-.01,.01)),   # w_{cue_2,t} w_{cue_2-opt_1,t} w_{cue_2-opt_2,t}
  lambda = .2,
  flo = -1,
  ceil = 1
){
  # Equation (4)
  delta =
    lambda * rowSums(t((desiredActiv-activ) *
                      weightsCuesOptions[2:3,]))
  # Equation (5) and Equation (6)
  updatedWeights =
    weightsCuesOptions[1,] +
    delta * ifelse(round(weightsCuesOptions[1,], 20) < 0,
                   weightsCuesOptions[1,] - flo,
                   ceil - weightsCuesOptions[1,])
  # prompt output of function
  return(updatedWeights)                  # w_{cue_1,t+1} w_{cue_2,t+1}
}
```

Note. Comments are in gray and R-commands are in bold.

D. Differences in predictions of decision times between both PCS-implementations

Predictions for both implementations of PCS are highly intercorrelated with $r = .81$ ($p < .001$). A scatterplot of iterations for all participants and trials (i.e., 69,525 data points) between both models reveal that $\text{PCS}_{\text{Transf}}$ and $\text{PCS}_{\text{Noise}}$ tend to differ only for iterations above 175 (Figure 8, panel A). $\text{PCS}_{\text{Noise}}$ tends to predict the same decision time for trials in which $\text{PCS}_{\text{Transf}}$ still discriminates in its predictions (i.e., for $I_{Transf} \gtrsim 175$). A scatterplot between predictions of $\text{PCS}_{\text{Transf}}$ and decision time (Figure 8, panel B, black dots) and $\text{PCS}_{\text{Noise}}$ and decision time in the same plot (red dots) reveals that decision trials with iterations above 175 do not differ much in decision times as predicted by $\text{PCS}_{\text{Noise}}$. That is, each dot of the purple line gives the median of the observed decision times for tasks with predicted iterations for $\text{PCS}_{\text{Transf}}$ in a range of +/-5 iterations around the dot and each dot of the yellow line gives the median decision times for the same range for $\text{PCS}_{\text{Noise}}$. Look for example at the two dots at iteration 100 (they overlap for both lines). The purple dot is the median of the decision times for all black dots within the boundaries of the dotted vertical blue lines (i.e., all trials within this range for $\text{PCS}_{\text{Transf}}$). The yellow dot is the the median of decision times for all red dots within this range (i.e., all trials within this range for $\text{PCS}_{\text{Noise}}$). The medians of decision times increase as predicted by both models. Additionally, the medians of the purple line reach a plateau near the boundary of 175 iterations where PCS_{Noise} does not make different predictions between trials (i.e., most of the black dots beyond an iteration of 175 lie at the border of 175 for PCS_{Noise} as can be seen from panel A). Thus, the observed data-pattern supports the predictions of PCS_{Noise}. One potential post-hoc explanation for this pattern is that trials that result in higher iterations for PCS tend to be more difficult tasks and that at some level of difficulty decision times do not differ any longer (at least for this type of task). This is supported by the data: Decision trials with iterations above 175 tend to be less often solved correctly than decision trials with iterations below 175 with 68% versus 89% correct decisions.

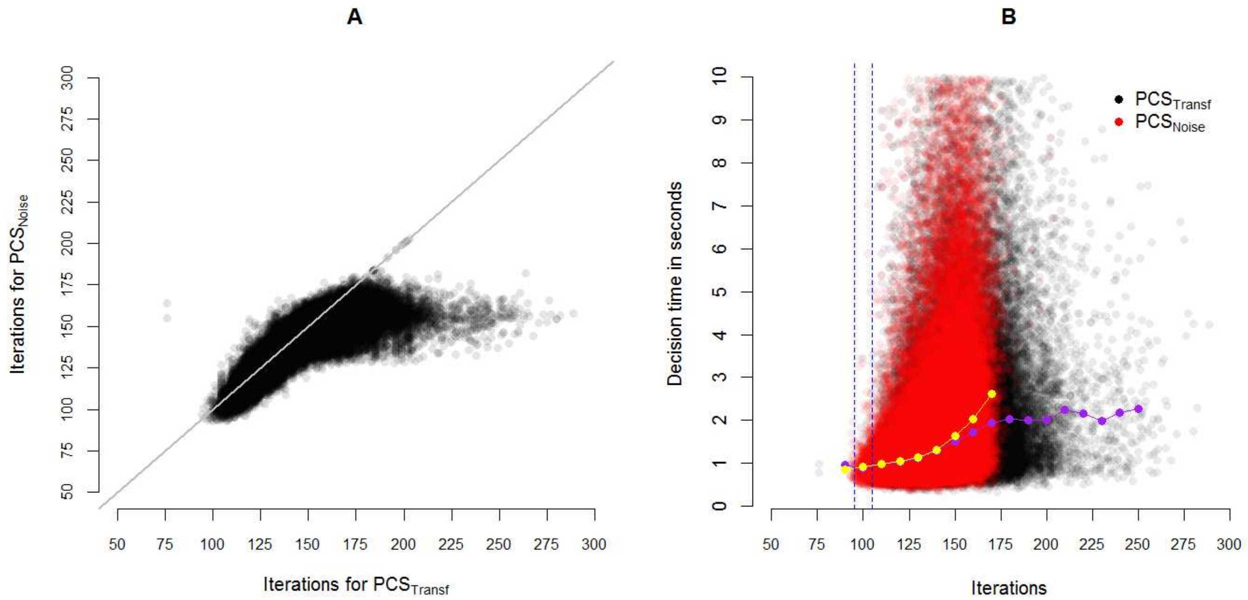

Figure 8. Scatter-plot between PCS-iterations of the decision process for all participants and tasks between PCS_{Transf} and PCS_{Noise} (A) and plot between iterations for both models and observed decision times up to ten seconds (i.e., $\sim$ 99% of all data) (B). The gray line in panel A indicates identical predictions (i.e., $y = x$) for both PCS implementations; for an explanation of the purple and yellow line in panel B see text.

E. Cross-validation

To test for the stability of results in relation to the method used for model-classifications (e.g., for avoiding potential problems of over-fitting of models; Marewski & Olsson, 2009; Myung, Pitt, & Kim, 2005), we re-ran the model classifications in all studies. In difference to the analyses reported

Table E.5. Percentage of participants best explained by model with free model-parameters fitted in the first round (study 1 and 2) or the first 50 trials (study 3) in each study and summed over studies based on choices in the final two rounds (study 1 and 2) or the final 50 trials (study 3) with fixed model-parameters.

	PCS_{Transf}	PCS_{Noise}	SSL	TTB	WADD	RAT
Study 1	8	81	1	0	9	0
Study 2	15	75	1	0	8	1
Study 3	30	40	17	2	7	5
Overall	17	68	5	0	8	2

in the main text, we fitted free model-parameters in the first 121 trials of the first round in study 1 and study 2 and the first half (i.e., fifty trials) in study 3 and based our classification on the remaining 121 trials $\times$ 2 final rounds $= 242$ trials in study 1 and 2 and the remaining fifty trials in study 3 with parameters fixed. Classifications lead to similar conclusions (Table E.5): PCS_{Noise} can explain the majority of participants (between 40% to 81% of participants) and both PCS-implementations can predict participants' decision best (between 70% to 90% of participants). In contrast to the results reported in the main text, SSL performs slightly better in the third study with 17% of all participants explained best by SSL. Summed over all three studies, percentages of classified participants for each model are similar (compare with Table 3) except for a minority of 8% of participants that can be best explained by the single-strategy WADD.