Transfer of learned category-response associations is modulated by instruction

Cai S. Longman[1], Fraser Milton[1], Andy J. Wills[2] and Frederick Verbruggen[1,3]

[1] University of Exeter, Exeter, UK

[2] Plymouth University, Plymouth, UK

[3] Ghent University, Ghent, Belgium

Abstract: 234 words. Text: 17714 words, 74 references, 9 figures, 6 tables.

Running head: Instructed category-response associations

Corresponding author:
Cai S. Longman
Psychology, CLES
University of Exeter
Washington Singer Building
Perry Road
Exeter EX4 4QG
United Kingdom
e-mail: c.s.longman@exeter.ac.uk
tel: +44 (0)1392 724626
fax: +44 (0)1392 724623

Funding: This work was supported by the European Research Council (grant number 312445). Frederick Verbruggen is a Royal Society Wolfson Research Merit Award holder.

Acta Psychologica (in press). Doi: 10.1016/j.actpsy.2017.04.004

Abstract

Although instructions often emphasize categories (e.g., odd number → left hand response) rather than specific stimuli (e.g., 3 → left hand response), learning is often interpreted in terms of stimulus-response (S-R) bindings or, less frequently, stimulus-classification (S-C) bindings with little attention being paid to the importance of category-response (C-R) bindings. In a training-transfer paradigm designed to investigate the early stages of category learning, participants were required to classify stimuli according to the category templates presented prior to each block (Experiments 1-4). In some transfer blocks the stimuli, categories and/or responses could be novel or repeated from the preceding training phase. Learning was assessed by comparing the transfer-training performance difference across conditions. Participants were able to rapidly transfer C-R associations to novel stimuli but evidence of S-C transfer was much weaker and S-R transfer was largely limited to conditions where the stimulus was classified under the same category. Thus, even though there was some evidence that learned S-R and S-C associations contributed to performance, learned C-R associations seemed to play a much more important role. In a final experiment (Experiment 5) the stimuli themselves were presented prior to each block, and the instructions did not mention the category structure. In this experiment, the evidence for S-R learning outweighed the evidence for C-R learning, indicating the importance of instructions in learning. The implications for these findings to the learning, cognitive control, and automaticity literatures are discussed.

Keywords: instructed learning, S-R learning, automaticity, cognitive control, categorization

A remarkable feature of human performance is the ability to rapidly learn and perform novel tasks from simple instructions. Instructions usually specify particular stimulus-response (S-R) mappings (e.g., X → left index finger, O → right index finger in a simple two-choice task) or slightly more complex/abstract category-response (C-R) mappings (e.g., odd → left hand, even → right hand in a digit classification task; living → index finger, non-living → middle finger in an object classification task). According to Chein and Schneider's (2012) triarchic theory of learning, a *metacognitive* system allows the rapid acquisition of such mappings by orchestrating (and then monitoring) the activity of a *cognitive control* system during the very early stages of learning. This is achieved by initiating (and terminating) the control routines that make successful initial performance possible and then monitoring their progress in order to enhance performance/learning by modifying any unsuccessful routines. The cognitive control system remains active throughout learning (under the guidance of the metacognitive system) and monitors, organizes, and alters the activity of a lower-level *representation* (associative learning) system to maximize efficient performance. More specifically, during the early stages of learning the cognitive control system is responsible for directing attention toward task-relevant information and away from distractions according to the current task goals. It is also responsible for selecting, updating and sequencing task-relevant actions and, as learning progresses, adjusting task parameters following suboptimal outcomes (under the direction of the metacognitive system). Thus, the early stages of learning can be characterized as the orchestration and monitoring of information processing towards the current goal and is largely under the guidance of the metacognitive and cognitive control systems.

After sufficient practice, performance is mostly supported by the representation system (Chein & Schneider, 2012). Theories of automaticity assume that performance has become automatized when exposure to a stimulus directly elicits an associated response. Schneider and Shiffrin (1977) distinguished between consistent and varied mappings of stimuli onto responses. In *consistent mapping*, the stimulus is consistently mapped onto the same response throughout practice, whereas in *varied mapping*, the stimulus is inconsistently mapped onto different responses throughout practice. In consistent mapping, associations between the stimulus and response are formed and automatic processing develops across practice. In varied mapping, inconsistent stimulus-response associations are formed, thereby preventing automatic processing (Shneider & Shiffrin, 1977). In a similar vein, Instance Theory (Logan, 1988, 1990) construed automaticity as a memory phenomenon. Initially, people would perform a task based on task rules (algorithmic processing). But after every stimulus encounter, they would store a new processing episode, which consists of a specific combination of the stimulus, the interpretation given to the stimulus, the response, and the task goal. When the stimulus is repeated, previous processing episodes are retrieved, facilitating performance when the mapping is consistent, but impairing performance when the mapping is inconsistent. Eventually, in consistent-mapping conditions, performance can rely entirely on memory retrieval (bypassing the cognitive control system) and is said to be ‘automatic’.

Most work on learning and automatization has focused on the formation of specific associations between stimuli and responses (Hazeltine & Schumacher, 2016). However, some research has questioned the way a ‘stimulus’ and ‘response’ should be conceptualized (e.g., Henson, Eckstein, Waszak, Frings, & Horner, 2014) and the relative importance of S-R associations (e.g., Hazeltine & Schumacher, 2016; Logan, 1990). For example, Horner and Henson (2009, 2011) asked participants to classify pictures of everyday items in a study-test design in which the stimulus (picture vs. word), the action (left vs. right button press), the decision (yes vs. no) or the classification (e.g., larger than a shoe box vs. larger than a wheelie bin) could change between the study and test phases. They found that at least two levels of stimulus representation (specific stimulus vs. abstract/semantic representation) could independently become associated with at least three levels of response representation (action, decision, classification). In a similar vein, Moutsopoulou, Yang, Desantis, and Waszak (2015; see also Moutsopoulou & Waszak, 2012, 2013; Waszak, Hommel, & Allport, 2004) have compared the formation and durability of stimulus-action and stimulus-category (S-C) associations. They also asked participants to classify pictures of everyday items and manipulated whether the classification and/or response (action) switched or repeated between prime and probe phases. Like Horner and Henson, Mousopoulou and colleagues confirmed that S-R and S-C associations are relatively independent (see also Dreisbach, 2012, for a review of recent research investigating the importance of task rules in modulating performance). Finally, Allenmark, Moutsopoulou, and Waszak (2015) have demonstrated that stimulus-action and stimulus-category associations do not depend on very low-level perceptual features (e.g., color), which led them to conclude that higher level representations (e.g., objects or semantic classifications) become associated with categories or actions (see also Frings, Moeller, & Rothermund, 2013, and Denkinger & Koutstaal, 2009; but for an example of evidence to the contrary, see Schnyer et al., 2007). Combined, these studies indicate that people can learn different types of associations when they perform a task. Learning different types of associations might even be the norm (e.g., Dreisbach, 2012; Hall, 2002; Verbruggen, Best, Bowditch, Stevens, & McLaren, 2014). The research summarized above has investigated several such associations, but none has offered a direct comparison between C-R associations (independent of the stimulus), S-R associations (independent of the classification), and S-C associations (independent of the response). The aim of the present study was to compare the relative contribution of these types of associations to learning by assessing the extent to which they transferred to novel stimuli, classifications and responses (respectively). Of particular interest was the relative contribution of C-R associations to learning which has thus far been the subject of few experimental reports.

C-R associations in the control and learning literature

C-R associations (e.g., odd → left hand response) are presumably an important part of rule-based performance. But despite being regularly utilized when instructing people how to perform a task, C-R associations have received little attention in the automaticity and control literature. Where it has been investigated experimentally, research has largely been dominated by visual-search paradigms (e.g., Kramer, Strayer & Buckley, 1991; Neisser & Beller, 1965; Schneider & Fisk, 1984) and/or the use of well-learned taxonomic categories such as letters, numbers, animals, colors, etc. (e.g., Neisser & Beller, 1965; Pashler & Baylis, 1991; Schneider & Fisk, 1984). Although these reports have been informative and largely indicate that learned C-R associations transfer to novel stimuli from the practiced categories, it is not possible to generalize their findings to more abstract category structures. Kramer, Strayer and Buckley (1990) note two particularly relevant reasons why the use of well-learned taxonomic categories is not ideal in this regard: (1) it is possible that a portion of the observed transfer effect could be due to extra-category associations (e.g., 'cat' and 'dog' might be associated by the phrase 'raining cats and dogs') rather than the experimenter-defined category structure (e.g., 'animals'); (2) it is possible that the observed transfer effect is limited to those members of the category that have been learned prior to the experimental session and does not generalize to novel exemplars that adhere to the category rules but were not known prior to testing (e.g., 'cat' and 'dog' are well-known animals that are likely to benefit from transfer, but 'caracal'[1] is less well-known and is therefore less likely to benefit from transfer in experiments that use word stimuli despite also being a member of the category 'animals'). Thus, a critical part of (instructed) learning is the ability to rapidly apply novel rules, but the use of well-learned taxonomic categories in research investigating C-R associations necessarily limits the extent to which the results can be generalized.

[1] A caracal is a rare wild cat that lives in Africa.

In an attempt to address the above criticisms, Kramer et al. (1990) used 'artificial' rule-based categories in two experiments investigating the development and transfer of automatic processes. The target stimuli were four concentric circles with two digits presented at random locations within their boundaries and the task was to determine whether the digit values and locations were consistent with rule-defined categories such as "1 ring apart, outer = inner" (i.e., are the digits presented one ring apart and are their values equal?). Kramer and colleagues observed effective transfer of learned C-R associations to novel exemplar stimuli, which is consistent with the notion that C-R associations make an important contribution to learning. However, their design does not allow for a clean measure of S-R learning independent of the classification because each stimulus could only be consistent with a single category (rule). Thus, a change of S-R mapping was necessarily confounded by a change of classification, making a direct comparison between S-R and C-R associations impossible.

More recently, Cohen-Kdoshay and Meiran (2007, 2009) have demonstrated that the flanker congruency effect (i.e., better performance on trials in which irrelevant stimuli presented alongside the target stimulus afford the same response as the target relative to trials on which the irrelevant stimuli afford a different response to the target) can be observed on the first trial following some simple C-R instructions (e.g., letter from the first half of the alphabet → left hand response, letter from the last half of the alphabet → right hand response). Cohen-Kdoshay and Meiran framed their discussion in terms of S-R bindings, but their results suggest that all members of the instructed categories automatically activate the relevant responses, even when they should be ignored. However, because all of the relevant exemplar stimuli used in the subsequent block were presented during the instructions phase, it is not possible to make any strong claims regarding the extent to which participants activated C-R bindings independent of the specific stimuli they were presented with.

Cohen-Kdoshay and Meiran's experiments have also initiated a recent interest in another relevant line of research investigating intention-based reflexivity – reflexive performance of an instructed S-R mapping even if the context is not appropriate (e.g., Meiran, Pereg, Kessler, Cole, & Braver, 2015a, 2015b; Liefooghe, De Houwer, & Wenke, 2013; Liefooghe, Wenke, & De Houwer, 2012). For example, Liefooghe Wenke and De Houwer (2012) instructed participants in two S-R mappings (the 'inducer' task) which only became relevant after completion of a second ('diagnostic') task which used the same stimuli as the inducer task but their identity was irrelevant (the diagnostic task was to classify the stimuli based on their orientation). By comparing performance on trials in the diagnostic task where the required response was either compatible or incompatible with the responses for the inducer task Liefooghe and colleagues were able to demonstrate that performance is modulated by instructed S-R bindings even if the context is not relevant to the instructed task and before the instructions have been carried out (Meiran et al, 2015a, 2015b, even found the effect on the first trial following the instructions). Interestingly, Liefooghe and colleagues also manipulated the specific effectors/modality used to respond in each task (index fingers, middle fingers, vocal response) and found clear evidence of intention-based response congruency (better performance on diagnostic task trials in which the required response was the same as the equivalent response for the inducer task relative to when the required response was different) even when the tasks used different effectors (index fingers vs. middle fingers) or response modalities (manual response vs. vocal response). The latter finding suggests that the representation of the instructed S-R bindings is relatively abstract and include the concepts 'left' and 'right' rather than specific actions.

Thus, of the few studies that focus on (or include) C-R effects, many used well-learned taxonomic categories and/or were not able to directly compare clean measures of S-R learning and C-R learning. The current experiments were designed to directly investigate the extent to which instructions can influence how novel rules are represented and to allow the direct comparison of S-R vs. C-R associations by independently manipulating repetitions of the stimulus, classification and/or response in a training-transfer design (details provided below). Finally, the research investigating intention-based reflexivity was expressly designed to investigate the formation of S-R (or C-R) associations through instruction alone (i.e., they did not investigate learning beyond the instructions phase), whereas the current experiments focused on the early stages of learning by investigating which associations transferred when one or more task components changed.

C-R associations in the categorization literature

C-R associations have also received some attention in the categorization literature (e.g., Ashby, Ell, & Waldron, 2003; Kruschke, 1996; Maddox, Glass, O'Brien, Filoteo, & Ashby, 2010; Nosofsky, Stanton, & Zaki, 2005; Wills, Noury, Moberly, & Newport, 2006). Most categorization studies have examined how stimuli are categorized and how categories are represented (for some reviews see Murphy, 2002; Pothos & Wills, 2011; Wills, 2013). Two theories have dominated the categorization field for a number of years. Prototype theories (e.g., Homa, Cross, Cornell, Goldman, & Shwartz, 1973; Reed, 1972; Smith & Minda, 1998, 2002) assume that novel stimuli are categorized by comparison to a prototype (a mental representation of the commonalities between category members or an idealized exemplar) from all likely categories. On the other hand, exemplar theories (e.g., Medin & Schaffer, 1978; Nosofsky, 1986; Nosofsky & Palmeri, 1997) assume that novel stimuli are compared to existing exemplars from all likely categories and assignment is based on similarity to the exemplars[2]. Although these theories are commonly contrasted with each other, many of their core ideas are shared. For instance, they both assume that categorization relies on

[2] It should be noted that some of Nosofsky's later work was, at least in part, inspired by Logan's (1988) Instance Theory of automatization. Indeed, Nosofsky and Palmeri's (1997) exemplar-based random walk model of speeded classification combines elements of Nosofsky's (1986) generalized context model of categorization and Logan's (1988) Instance Theory of automatization.

comparing novel stimuli to existing category information, and that membership is decided based on similarity. However, prototype theories must assume the existence of category-level representations whereas exemplar theories need not (although some do; see Kruschke, 1996).

Much categorization work has focused on the representation issue. In other words, it has mostly focused on whether the memorial representation of categories includes the storage of specific examples, prototypes, or both. In other words, it has largely focused on S-C associations (e.g., is stimulus A represented under category X) and less on C-R associations (e.g., category X → left index finger response). The literature on cognitive control and learning described above suggests that categories are important for guiding performance (e.g., instructions often map categories onto responses without mentioning all exemplars), but it remains largely agnostic regarding the nature of the category representation. Addressing the representation issue (which is still debated in the categorization literature) is beyond the scope of the current report, but we will consider it when interpreting the results from the experiments.

The present study

The present study consists of five experiments designed to investigate the relative contribution of C-R, S-R and S-C associations to instructed category learning. Experiments 1-3 were designed to compare C-R and S-R learning and to determine the extent to which each kind of association transfers beyond the specific learning context (i.e., whether C-R associations transfer to novel stimuli from the same category and/or whether S-R associations transfer across classifications). To preempt the results, we found some evidence of both C-R and S-R learning. However, the evidence that C-R associations readily transfer to novel stimuli was much stronger than the evidence for transfer of S-R associations across classifications. Indeed, we even found some evidence that S-R learning suffered from interference when the classification changed.

Experiment 4 extended this comparison to stimulus-category (S-C) associations and found that, in the current paradigm, there was a bias toward C-R learning that outweighed the effect of both S-R and S-C learning. Finally, Experiment 5 reduced the C-R bias in the instructions by mentioning (and presenting) only the S-R bindings. This manipulation was enough to shift the performance bias away from C-R learning and toward S-R learning.

Experiment 1

We devised a task where individual perceptual stimuli (dot patterns) were classified according to their category membership by entering a 4-digit response 'code' (we used complex responses to guarantee that novel responses could be used in each task). Participants in categorization experiments are typically required to learn the categories over long sequences of trials accompanied by feedback, but without any formal instructions regarding the categories themselves (e.g., Wills et al., 2006; though for a notable exception see Allen & Brooks, 1991). In a similar vein, the automaticity experiments designed to investigate C-R associations described above required participants to learn the relevant C-R bindings over long training sessions (sometimes lasting several days, or even weeks; e.g., Kramer et al., 1990, 1991; Neisser & Beller, 1965). In the current experiments we were interested in the very early stages of instructed category learning so we presented participants with the main surface features of the task at the start of each block – i.e., participants were presented with the category templates (all templates can be found in the Appendix) and the correct response code for each category. We were thus able to examine what kinds of associations are learned over the first few trials of instructed (perceptual) category learning tasks that use novel stimuli and categories[3].

[3] We owe much of our experimental design to the recent interest in the investigation of rapid instructed task learning (e.g., Cole, Bagic, Kass, & Schneider, 2010; Ruge & Wolfensteller, 2010). In these studies participants perform multiple tasks that share a common structure but differ in their surface features and the data are collapsed across tasks.

The experimental session was divided into pairs of ‘training’ and ‘transfer’ blocks. Entirely novel stimuli, category templates and response codes were introduced at the beginning of each training block. Each training block was immediately followed by a transfer block. In half of the transfer blocks, the categories were repeated from the preceding training block; in the other half, novel categories were introduced. Whether the classifications changed or not, half of the transfer stimuli were novel and half were repeated from the preceding training block (i.e., stimulus and category repetitions were manipulated independently by using stimuli that equally resembled two category templates; see Figure 1 for an overview of the design and some example stimuli from each condition). In other words, we used a two (Stimulus: same vs. different) by two (Category: same vs. different) design. In one critical condition the stimuli were novel, but the categories and responses were repeated (i.e., the transfer phase used Different stimuli, but the Same categories and the Same responses as the preceding training phase; DsScSr[4]); this condition can be used as a clean measure of C-R learning (independent of S-R learning). In a second critical condition novel category templates were introduced at transfer, but the stimuli were repeated and they were mapped on the same responses as those used during training (SsDcSr); this condition can be used as a clean measure of S-R learning (independent of the classification). In the DsDcSr condition, both the stimuli and categories used at transfer differed from the ones used in training; this will be our baseline condition. Finally, everything remained the same in the SsScSr condition; thus, this condition included both S-R and C-R learning.

We chose to manipulate category repetitions across blocks and stimulus repetitions within blocks primarily to ensure that the categories played an important role in the SsDcSr condition and to potentially allow C-R transfer in the DsScSr condition. If the SsDcSr condition was in a block of its own (without the DsDcSr condition) then it might have been much easier for participants to

[4] The capital letters (and lower-case letters) indicate whether the stimuli (s), categories (c) and/or responses (r) of the transfer phase were the same (S) or different (D) from those used in the training phase.

ignore the introduction of novel category templates at transfer and instead rely entirely on S-R learning. If the DsScSr condition was in a block of its own (without the SsScSr condition) then participants might have been less likely to consider the novel stimuli as part of an existing category, despite the instructions. Thus, to ensure that participants processed the instructions presented at transfer, we used a mixed-block manipulation. Note that this manipulation also allowed more training-test pairs, increasing the number of observations and optimizing the signal to noise ratio.

We could determine what was learned during training by comparing initial transfer performance from the various conditions. If participants were able to transfer learned S-R associations to novel classifications (in the absence of C-R learning) then initial performance at transfer should be comparable to performance at the end of training in conditions where the S-R bindings are repeated, whether the classifications were repeated or not (i.e., initial transfer performance for SsScSr ≈ SsDcSr < DsScSr ≈ DsDcSr). If participants were able to transfer learned C-R associations to novel stimuli from the same category (in the absence of S-R learning) then initial transfer performance should be comparable to performance at the end of training for conditions where the C-R bindings were repeated, whether the stimuli were repeated or not (i.e., initial transfer performance for SsScSr ≈ DsScSr < SsDcSr ≈ DsDcSr). If both S-R and C-R learning contributed to performance, then initial transfer performance should be comparable to performance at the end of training in the SsScSr condition, intermediate in the DsScSr and SsDcSr conditions, and worst in the DsDcSr condition (i.e. initial transfer performance for SsScSr < DsScSr ≈ SsDcSr < DsDcSr). Finally, performance in all conditions should be comparable to baseline if subjects could not transfer S-R or C-R associations (i.e. SsScSr ≈ DsScSr ≈ SsDcSr ≈ DsDcSr).

Method

Participants. 40 students from the University of Exeter (37 female) with a mean age of 19.6 years (SD = 2.0) participated for £7 or partial course credits. The target sample size and exclusion criteria were decided in advance of data collection (when N = 40, we could detect medium-sized differences). All experiments of the present study were approved by the local research ethics committee at the School of Psychology, University of Exeter. Written informed consent was obtained after the nature and possible consequences of the studies were explained.

Apparatus, stimuli, and responses. Stimuli were presented on a 21.5-inch iMac using Psychtoolbox (Brainard, 1997). As shown in Figure 2, the stimuli consisted of patterns of five black dots (diameter = 0.5 cm) presented at a pseudo-random location in a larger array (18 x 18) of small gray dots (diameter = 0.25cm, distance between adjacent dots = 0.75 cm), which was itself surrounded by a black square (side = 15 cm, thickness = 0.1 cm). Randomizing the location of each stimulus ensured that the stimuli could not be classified according to their relative location on the screen, and therefore, encouraged comparison with the category templates.

In each block, there were two categories (so two category templates, selected pseudo-randomly), and eight stimuli (four exemplar stimuli per category). The category templates are shown in the Appendix. Each stimulus included five black dots. Three were positioned within the borders of the template and the remaining two were presented at adjacent locations (see Figure 1 for some examples). Thus, the category membership of a given dot-pattern was determined by its overall similarity to the category templates presented at the start of each block. On each trial the participant indicated to which category the stimulus belonged by entering the relevant four-digit response code with their preferred index finger (all participants used their right hand) using the numeric keypad on a standard keyboard.

Complex response codes were used so that novel responses could be associated with each category-pair presented throughout the experiment. The response codes all started and ended with '5' to ensure the index finger was in the optimal (central) position at the start of each trial. The intervening digits were always on adjacent keys (e.g., 5235, 5425) in order to equate the difficulty of entering each response code. Two different response codes were used for each training-transfer phase (one per category). The relevant codes used in each phase were selected pseudo-randomly to discourage re-classification according to a simple spatial rule and also to equate the difficulty of each phase (e.g., when used in the same block, the codes 5475 and 5415 could easily be reclassified as 'up' and 'down' respectively because the only difference between them is the digit following the 4 – either an 'upward' or 'downward' motion with the index finger; therefore, they might be easier to memorize than 5235 and 5425).

Procedure. At the beginning of the experiment, participants were informed that each block would start with an instruction screen (displaying the two category templates and relevant response codes for the block; see Figure 2) and end with a feedback screen (displaying their mean correct RT, number of errors, and proportion of correct responses). Both the instructions screen and feedback screen were displayed for 15 seconds each. Participants were instructed to respond as quickly as possible while minimizing errors. They were also informed that each response code would start and end with '5' so they should place their index finger on that key before the experiment started.

In each block, all trials started with a blank screen presented for 500 ms, followed by the target stimulus which was visible until the first digit of the response code was entered. There was no response deadline. The response code appeared on the screen as it was typed in. Immediate feedback, visible for 1000 ms, was presented after the final digit had been entered: the stimulus was displayed as it appeared during the trial with the category template presented behind (to confirm category membership) in either green (correct) or red (error), and the correct response code was presented directly below in black Arial font (size 30, see Figure 2).

Training and transfer blocks alternated throughout the experiment (i.e., training–transfer–training–transfer–…). Each block used two categories assigned to one response each. There were four exemplar stimuli per category and each stimulus was presented ten times, resulting in 80 trials per block. During the training blocks all stimuli, category templates and response codes were novel. In the transfer blocks, the category templates could either be novel or repeated from the preceding training block. Whether the category templates were repeated or not, half of the stimuli used during transfer were novel and the other half were repeated from the preceding training block (the latter stimuli equally belonged to two categories – one category was always used during training, the other category was only used in transfer blocks where novel categories were introduced; all category templates from the training and transfer blocks are presented in the Appendix). When the category templates were repeated between training and transfer, the novel stimuli introduced at transfer were also based on the category templates used during training. When novel category templates were introduced at transfer, the novel stimuli were based on the novel category templates shown at the beginning of the transfer block (see Figure 1 for some example stimuli from the training and transfer blocks in each condition). The responses (e.g., 5235, 5425) were always repeated between training and transfer blocks. This two (Stimulus: same vs. different) by two (Category: same vs. different) design resulted in four conditions defined on the basis of whether the stimuli and/or categories[5] used in the transfer block were the same or different to those used in the preceding training block: Different stimuli, Different categories, Same responses (DsDcSr); Same stimuli, Different categories, Same responses (SsDcSr); Different stimuli, Same categories, Same responses (DsScSr); Same stimuli, Same categories, Same responses (SsScSr). Note that each

[5] Although the responses never changed between training and transfer in Experiments 1, 2, 3, and 5, response repetition was manipulated in Experiment 4. For consistency the condition codes indicate whether the responses used at transfer where the same or different to those used during training for all experiments.

transfer block consisted of two conditions: transfer blocks in which the categories changed, consisted of DsDcSr and SsDcSr trials; whereas transfer blocks in which the categories were repeated consisted of DsScSr and SsScSr trials. Whether the category templates used at transfer were the same or different to those used during training alternated through the experiment and the order was counterbalanced over participants.

The experiment started with a practice session, which consisted of one training-transfer pair (80 trials per block). These blocks were omitted from all analyses. The experimental session consisted of 12 blocks of 80 trials each (3 training-transfer pairs where the categories changed, and 3 training-transfer pairs where the categories were repeated). This resulted in a total of 120 training trials and 120 transfer trials available for analysis per condition (e.g., there were three transfer blocks in which the categories changed, and in each of these blocks, there were 40 DsDcSr trials and 40 SsDcSr trials).

The participants were not informed that the first two blocks would be omitted from analysis or that the experiment used a training-transfer design. In addition to the 15-second breaks at the beginning (instruction screen) and end (feedback screen) of each block, a timed 2-minute break was enforced after block number 8 (half way through the experimental session).

Analyses. All data processing and analyses were performed using R (R Development Core Team, 2015). Raw data files and R scripts from all experiments are deposited on Dropbox (https://www.dropbox.com/sh/7v4339cjvo3ydla/AAC3pMcn78T5PJ-7noA_vTDQa?dl=0; if this paper is accepted the data files and R scripts will be deposited on the Open Research Exeter data repository).

Trials with RT (first digit) <100 ms (0.05%) and trials with RT (last digit) >5000 ms (0.78%) were omitted from all analyses. Error trials were omitted from the RT analyses. Because the first digit in each response code was the same, it was possible for participants to execute the first digit before the stimulus was categorized. A more informative measure of performance was the latency of the final digit in the response code (i.e., the time at which the full response had been entered), so we used this for all response latency analyses.

Performance from both categories used in each condition was pooled. Transfer performance was compared to training performance for the corresponding stimuli: for the transfer conditions in which the stimuli were repeated (i.e., SsScSr and SsDcSr), we compared transfer performance to the average training performance from the stimuli that were repeated at transfer; for the transfer conditions in which the stimuli changed (i.e., DsScSr and DsDcSr), we compared transfer performance for the novel stimuli with the average training performance for the stimuli that were replaced at transfer (see Figure 1). For the statistical analyses, we calculated three measures in each condition for every participant (see Figure 3; these measures are based on the measures used in Logan, 1990): (a) *training learning* (the averaged performance from the first three presentations of the relevant stimuli for the condition during the training blocks minus the averaged performance during the final three presentations of the relevant stimuli for the condition during the training blocks[6]); (b) *transfer effect* (the averaged performance from the first three presentations of the relevant stimuli for the condition during the transfer blocks minus the averaged performance from the final three presentations of the relevant stimuli for the condition during the training blocks); and (c) *transfer learning* (the averaged performance from the first three presentations of the relevant stimuli for the condition during the transfer blocks minus the averaged performance from the final three presentations of the relevant stimuli for the condition during the transfer blocks).[7] The first and last of these measures quantify overall learning during the training and transfer blocks

[6] Mean correct RT was calculated by averaging the performance from the first/last three stimulus presentations in a given block after errors from these trials had been omitted.

[7] A potentially more informative method for such an analysis would be to perform nonlinear regression. We did perform such analyses on data from a pilot study in which we attempted to fit a power function and an exponential function (Logan, 1988) to the learning data. However, some of the conditions had quite flat learning curves which resulted in some extreme values being generated. For this reason we settled for the simpler analyses reported here.

respectively. Any notable differences between conditions in the ‘training learning’ measure would suggest some intrinsic difference between the rates of learning in the various conditions; this would complicate the interpretation of both transfer measures. Our main measure of interest is the ‘transfer effect’ as it can show what was learned during training. For example, if participants were able to transfer learned C-R associations to novel stimuli then the ‘transfer effect’ in the DsScSr condition should be relatively small in comparison to the baseline (DsDcSr) condition. On the other hand, if participants were able to transfer learned S-R associations across classifications then the ‘transfer effect’ should be relatively small in the SsDcSr condition in comparison to baseline. Finally, for completeness, we also analyzed the ‘transfer learning’ measure. It could mirror the ‘transfer effect’ measure, as there is more opportunity to learn when the ‘transfer effect’ is also large. However, it is possible that the introduction of novel task elements at transfer might hinder learning as well as initial transfer performance (i.e. large ‘transfer effect’ and no learning in the transfer block).

For each score and dependent variable (RT and error rate), we performed a separate Stimulus (same, different) by Category (same, different) ANOVA. We also report Bayes factors and effect sizes (generalized eta squared) for all relevant effects/interactions. Bayes factors were calculated with the BayesFactor package for the R Software environment (R Development Core Team, 2015), using the default JZS prior (.707; Morey, Rouder, & Jamil, 2015). To reduce the number of model comparisons, interactions were only allowed if all constituent sub-effects were also included (see Morey, Rouder, & Jamil, 2015). When this approach is used, Bayes factors less than 1 indicate that removing any effect/interaction from the full model is deleterious (i.e., a Bayes factor <1 indicates that the effect/interaction is a contributor to the fit of the full model). One-sample t-tests were performed on the ‘training learning’ measure to confirm that learning had taken place in each condition during the training blocks. For visualization purposes, we used ‘rolling means' to smooth the curves – each point in Figure 4 included the data from three stimulus repetitions, and the windows overlapped.

Results and Discussion

Mean RTs and the proportion of errors made in the training and transfer blocks are plotted as a function of condition (DsDcSr, DsScSr, SsDcSr, SsScSr) and stimulus repetition in Figure 4. The mean 'transfer effect' for RTs and the proportion of errors are plotted as a function of condition in Figure 5. The results from the omnibus ANOVAs are reported in Table 1.

'Training learning'. As expected, none of the effects or interactions were reliable in the ANOVAs on the 'training learning' measure for either RTs (Fs<1; BFs>3.4) or the proportion of errors (Fs<1; BFs>3.5) indicating that learning was comparable in all conditions during training. The one-sample t-tests on the 'training learning' measure found that learning had taken place in all conditions during the training blocks (RTs: ts>8.4; accuracy: ts>4.7).

'Transfer effect'. The ANOVAs on the 'transfer effect' for RTs and the proportion of errors found that the decrement in performance between the end of training and the start of transfer was much larger when novel categories were introduced at transfer (RT difference = 133 ms; accuracy difference = 6.1%) relative to when the categories were repeated (RT difference = 4 ms; accuracy difference = 0.6%; main effect of Category for RTs: F=46.18, BF<0.001 ± 5.7%; main effect of Category for accuracy: F=18.17, BF<0.001 ± 20.0%). Neither the main effect of Stimulus nor the Stimulus by Category interaction reached significance for RTs (Fs<3.3, BFs>2.5) or the proportion of errors (Fs<2.0, BFs>1.7) indicating that, although initial transfer performance was strongly modulated by whether the categories were novel or repeated from training, initial transfer performance was not materially affected by the introduction of novel stimuli.

'Transfer learning'. The ANOVAs on the 'transfer learning' measure found that the improvement in performance through the transfer block was much larger when novel categories were introduced at transfer (RT improvement = 129 ms; accuracy improvement = 5.8%) relative to when the categories were repeated (RT improvement = 37 ms; accuracy improvement = 1.6%; main effect of Category for RTs: F=14.30, BF<0.001 ± 14.0%; main effect of Category for accuracy: F=9.96, BF=0.003 ± 4.9%). Neither the main effect of Stimulus nor the Stimulus by Category interaction reached significance for RTs (Fs<1.4, BFs>2.6) or the proportion of errors (Fs<1, BFs>3.1) indicating that, although the amount learned during the transfer blocks was strongly modulated by whether the categories were novel or repeated from training, transfer learning was largely unaffected by the introduction of novel stimuli.

Taken together, this pattern of results confirms our hypothesis that, in the current paradigm, learned C-R associations readily transfer to novel stimuli from the same category but learned S-R associations do not transfer across classifications quite so easily.

Experiment 2

The results from Experiment 1 indicate that participants are able to rapidly learn C-R associations and transfer them to novel stimuli, whereas the evidence for transfer of S-R associations across classifications is sparse. Experiment 2 was designed with two goals in mind: first, we wanted to confirm the results reported in Experiment 1; second, we wanted to confirm our interpretation of those results. If our interpretation is wrong and S-R associations do play an important role in the current design then increasing the number of exemplar stimuli per category should further increase the C-R learning bias because there are many more S-R associations to learn and exposure to each stimulus will be reduced. On the other hand, if our interpretation is correct and C-R associations drive performance, it should follow that increasing the number of exemplars per category will not materially alter the pattern of results because the number of category repetitions will remain the same (maintaining the strength of the C-R bindings). In order to test this conjecture in Experiment 2, we included the same transfer conditions as in Experiment 1, but each condition was performed twice – once with 4 exemplar stimuli per category (as in Experiment 1) and once with 16.

Method

40 different students from the University of Exeter (36 female) with a mean age of 19.8 years (SD = 1.5) participated for the same reimbursement as Experiment 1 (£7 or partial course credits). The target sample size and exclusion criteria were identical to Experiment 1 and written informed consent was obtained after the nature and possible consequences of the studies were explained.

The apparatus, stimuli, responses and procedure were identical to Experiment 1 apart from the following critical difference: each condition from Experiment 1 was performed twice – once (as in Experiment 1) with four exemplar stimuli per category and once with 16 exemplar stimuli per category. The order of conditions was Latin-square balanced over participants.

The current experiment also differed from Experiment 1 on a number of minor structural details. Each block was divided into two 64-trial miniblocks with a break between each miniblock (the break followed the same format as the break between blocks in Experiment 1: 15 seconds feedback followed by 15 seconds instructions). The practice session prior to the experimental session consisted of only one training block (two 64-trial miniblocks). There were 8 further experimental blocks (4 training, 4 transfer): two training blocks were followed by transfer blocks in which the categories were repeated (DsDcSr and SsDcSr conditions; one with 4 exemplar stimuli per category, one with 16 exemplar stimuli per category); two training blocks were followed by transfer blocks in which novel categories were introduced (DsScSr and SsScSr conditions; one with 4 exemplar stimuli per condition, one with 16 exemplar stimuli per condition. A total of 18 miniblocks (including the practice session) of 64 trials each made up the entire experiment (total number of trials = 1152). Because of these changes to the structure of the experiment, the 2-minute break was after block number 10 (half way through the experimental session).

As in Experiment 1, trials with RT (first digit) <100 ms (0.23%) and trials with RT (last digit) >5000 ms (0.59%) were omitted from all analyses. Error trials were omitted from RT analyses. The data from three participants were replaced because they had <50% of the maximum possible observations in at least one condition (excluding the practice blocks) following the above data cleaning procedures (the same exclusion criterion was applied in Experiment 1 but it was not necessary to replace any participants). As in Experiment 1, we used rolling means for visualization of the results, but because the maximum number of stimulus repetitions differed between conditions (16 exemplars = 4 stimulus repetitions per block; 4 exemplars = 16 stimulus repetitions per block) the rolling windows were based on category repetitions (per condition within a given block) rather than stimulus repetitions (see Figure 6 for the windows). Finally, every condition was performed only once by each participant.

For each score and dependent variable (RT and error rate), we performed a separate ANOVA with the factors Stimulus (same, different), Category (same, different), and Exemplars (4, 16). We also report Bayes factors and effect sizes (generalized eta squared) for all relevant effects/interactions (calculated in the same way as for Experiment 1). One-sample t-tests were performed on the ‘training learning’ measure to confirm that learning had taken place in each condition during the training blocks. We also performed several follow-up paired-samples t-tests in order to unpack the significant interactions from the omnibus ANOVAs. We also report Bayes factors and effect sizes (Hedges’s average g (g_{av}); Lakens, 2013) for these comparisons.

Results and Discussion

Mean RTs and the proportion of errors made in the training and transfer blocks (separately for conditions where 4 or 16 exemplar stimuli were used per category) are plotted as a function of condition (DsDcSr, DsScSr, SsDcSr, SsScSr) and category repetition in Figure 6. The mean 'transfer effect' for RTs and the proportion of errors are plotted as a function of condition in Figure 5. The results from the omnibus ANOVAs are reported in Table 2.

'Training learning'. None of the effects or interactions were reliable in the ANOVA on the 'training learning' measure for RTs (*F*s<2.8; *BF*s>0.9). The ANOVA on the 'training learning' measure for the proportion of errors found that participants' response accuracy improved slightly more when there were 4 exemplar stimuli per category (improvement = 11.3%) than when there were 16 (improvement = 8.0%; main effect of Exemplars: ($F(1,39)=4.61$, $p=.038$, $\eta^2=.015$). However, the Bayesian analysis found only anecdotal[8] evidence that removal of the main effect of Exemplars from the model would impair its fit ($BF=0.954 \pm 8.7\%$). None of the other effects or interactions approached significance (*F*s<1.1, *BF*s>3.3). The one-sample t-tests on the 'training learning' measure found that learning had taken place in all conditions during the training blocks (RTs: *t*s>4.6; accuracy: *t*s>2.5).

'Transfer effect'. The ANOVAs on the 'transfer effect' for RTs and errors found that the decrement in performance between the end of training and the start of transfer was larger when novel category templates were introduced at transfer (RT difference = 82 ms; accuracy difference = 7.3%) relative to when the category templates were repeated (RT difference: 4 ms; accuracy difference: 1.0%; main effect of Category for RTs: $F=9.87$, $BF=0.016 \pm 17.5\%$; main effect of Category for accuracy: $F=18.44$, $BF<0.001 \pm 7.0\%$).

[8] We have adopted Wetzel et al.'s (2011) protocol for interpreting Bayes factors.

When novel category templates were introduced at transfer, the 'transfer effect' was smaller for altogether novel stimuli (DsDcSr RT difference = 50 ms; accuracy difference = 5.5%) than for those stimuli that had previously been classified under a different category (SsDcSr RT difference = 115 ms; accuracy difference = 9.0%). On the other hand, when the category templates were repeated between training and transfer, there was a decrement in initial transfer performance for novel stimuli (DsScSr RT difference = 45 ms; accuracy difference = 1.4%) but a slight improvement in RTs (SsScSr RT difference = -37 ms) and a smaller decrement in accuracy (SsScSr accuracy difference = 0.7%) for those stimuli that had been classified during training. This interaction was reliable for RTs (Stimulus by Category interaction: F=13.53, BF=0.028 ± 15.5%), but did not reach significance for accuracy (Stimulus by Category interaction: F=3.42, BF=1.914 ± 12.3%).

Follow-up paired samples t-tests found that the RT 'transfer effect' was 64 ms smaller in the DsDcSr condition relative to the SsDcSr condition ($t(39)$=2.41, p=.021, g_{av}=.355, BF=2.22) indicating that learned S-R associations suffered from considerable interference when they were performed under a novel classification (the 3.5% difference in accuracy between these conditions was not reliable: t=1.92, BF=0.90). Critically, the 'transfer effect' was significantly smaller in the DsScSr condition relative to the SsDcSr condition for accuracy (t=3.83, p<.001, g_{av}=.873, BF=62.61) but not for RTs (t=1.87, BF=0.83). The latter results suggest that the C-R bias in the RTs can largely be explained by interference for learned S-R associations when classified under novel categories, whereas the bias in accuracy can largely be explained by the rapid transfer of learned C-R associations to novel stimuli from the same category.

None of the other effects or interactions was reliable for RTs (Fs<1.3, BFs>3.2) or errors (Fs<3.0, BFs>1.5) indicating that, like in Experiment 1, initial transfer performance was modulated to a much greater extent by whether the transfer category templates were novel or repeated from training relative to the modest effect of introducing novel stimuli at transfer. Finally, the ‘transfer effect’ was comparable whether each category included 4 or 16 exemplar stimuli confirming our hypothesis that increasing the number of exemplar stimuli per category (and decreasing the exposure to each stimulus during training) had little effect on learning.

‘Transfer learning’. The ANOVAs on the ‘transfer learning’ measure for the proportion of errors found that the improvement in performance through the transfer block was much larger when novel categories were introduced at transfer (improvement = 7.0%) relative to when the categories were repeated (improvement = 1.8%; main effect of Category: F=12.24; BF=0.003 ± 14.4%). The improvements in RTs were also in the same direction (DsDcSr, SsDcSr average improvement = 130 ms; DsScSr, SsScSr average improvement = 80 ms), however the effect was not reliable (F=2.32; BF=0.624 ± 9.1%).

None of the other effects or interactions were reliable for RTs (Fs<3.2, BFs>2.4) or errors (Fs<3.2, BFs>0.8) indicating that, like in Experiment 1, transfer learning was modulated by whether the categories were novel or repeated from training, but not by the introduction of novel stimuli. Critically, the amount learned during transfer was comparable whether each category included 4 or 16 exemplar stimuli.

Summary. Experiment 2 was designed to confirm the results of Experiment 1 and to test whether increasing the number of exemplar stimuli per category would further increase the C-R learning bias reported in Experiment 1. We were able to confirm our hypothesis that learning was largely dictated by whether the categories used at transfer were novel or repeated from training, whereas the effect of introducing novel stimuli at transfer was comparatively small. Critically, none of the effects or interactions involving the factor Exemplars provided conclusive evidence of a systematic modulation of learning/performance by the number of exemplar stimuli per category. That dramatically increasing the number of exemplar stimuli per category did not materially change the pattern of results confirms our hypothesis that, at least in the current paradigm, C-R learning is a much stronger determinant of performance than S-R learning.

However, in the current experiment the evidence that learned C-R associations rapidly transferred to novel stimuli was relatively weak (by comparison to Experiment 1) but the introduction of novel category templates at transfer significantly hindered performance for learned S-R associations. The latter finding is consistent with the notion that performance is modulated by learned S-C associations (the interference to S-R learning is likely due to conflict as a result of familiar S-R bindings being performed under a novel classification), but does not necessarily support our hypothesis that learned C-R associations rapidly transfer to novel stimuli. The relative contribution of C-R, S-R and S-C associations to learning are directly compared in Experiment 4, but first we tested if the C-R findings from Experiments 1-2 could be due to our feedback protocol.

Experiment 3

In Experiments 1 and 2 the feedback given on every trial consisted of a highly salient image of the correct category template coupled with the correct response code. It could be argued that this protocol undermines our conclusions regarding C-R associations because on every trial the correct response is reliably coupled with a highly salient stimulus (i.e., the category template) that is perceptually similar to each of the exemplar stimuli from the same category. In comparison, the individual exemplar stimuli presented on each trial are relatively non-salient. Thus, what we describe as C-R associations could equally be classified as S-R associations because both the stimulus (the category template) and the relevant response are reliably paired during feedback. This could explain why we found that performance/learning was modulated by whether the categories were novel or repeated from training: when the categories were repeated the association between the response and the template is highly practiced whereas when novel categories are introduced at transfer the participant is required to learn novel S-R associations (i.e., an association between the new template and the response).

Experiment 3 was designed to rule out this possibility by altering the feedback procedure. To this end we assigned a meaningful name to each category (listed with the category templates in the Appendix) and the feedback consisted of presenting the correct category name and the correct response in either green (correct) or red (error). By removing the (perceptual) category template from the feedback display we eliminated any incidental formation of an S-R association between the category template and the response (though for an example of S-R effects when the stimulus switches between a picture and a word, or vice-versa, see Horner & Henson, 2011). It is possible that the category names activated a mental representation of the category template, but because the name did not perceptually resemble the exemplar stimuli (or the template) from the relevant category, a simple S-R account would be difficult to accommodate without relying on a more abstract representation of the category template (a verbal 'mediator' or a prototype).

Method

40 different students from the University of Exeter (29 female) with a mean age of 19.5 years (SD = 1.1) participated for the same reimbursement as in Experiments 1 and 2 (£7 or partial course credits). The target sample size and exclusion criteria were identical to Experiments 1 and 2, and written informed consent was obtained after the nature and possible consequences of the studies were explained.

The apparatus, stimuli, responses and procedure were identical to Experiment 1 apart from one critical change: each category was assigned a meaningful name (e.g., the categories presented in the training block of Figure 1 were named 'cross' and 'corner'; all category names are presented in the Appendix) so that feedback could be in the form of the correct category name and the correct response code presented in either green (correct) or red (error). The response code appeared in the center of the screen as it was entered by the participant (whereas in Experiment 1 it appeared below the stimulus area). Both the correct response code and the category name were presented centrally during feedback with the category name positioned one line above the response code. The category names were also presented alongside the templates and correct response codes on the instructions screen at the start of each block (see Figure 2).

As in Experiment 1, trials with RT (first digit) <100 ms (0.03%) and trials with RT (last digit) >5000 ms (1.20%) were omitted from all analyses, and error trials were omitted from RT analyses. One participant's data was replaced because they had <50% of the maximum possible observations in at least one condition (excluding the practice blocks) following the above data cleaning procedures. The analyses were identical to those performed in Experiment 1 except two additional paired-samples t-tests were performed in order to determine the relative contribution of C-R vs. S-R learning in the 'transfer effect' measure for the current experiment (these analyses were not necessary in Experiment 1 where the Stimulus by Category interaction did not approach significance in either of the omnibus ANOVAs). Bayes factors and effect sizes were also calculated in the same way as for Experiment 2 for these comparisons.

Results and Discussion

Mean RTs and the proportion of errors made in the training and transfer blocks are plotted as a function of condition (DsDcSr, DsScSr, SsDcSr, SsScSr) and stimulus repetition in Figure 7. The mean 'transfer effect' for RTs and the proportion of errors are plotted as a function of condition in Figure 5. The results from the omnibus ANOVAs are reported in Table 3.

'Training learning'. The ANOVA on the 'training learning' measure for RTs found that the improvement in performance through training was slightly larger for the stimuli that equally belonged to two categories (i.e., those that would subsequently be repeated at transfer whether the categories changed or not; SsDcSr, SsScSr average improvement = 308 ms) relative to the stimuli that belonged to a single category (i.e., those that would subsequently be replaced at transfer; DsDcSr, DsScSr average improvement = 264 ms; main effect of Stimulus: F=5.90). However, the Bayesian analysis found only anecdotal evidence that removal of the main effect of Stimulus from the model would impair its fit (BF=0.833 ± 10.4%). Neither the main effect of Category nor the Stimulus by Category interaction approached significance (Fs<1, BFs>4.0), nor did any of the effects or interactions in the ANOVA on the proportion of errors (Fs<1, BFs>4.0). Taken together these results indicate that the amount learned during training was broadly consistent across all four conditions. The one-sample t-tests on the 'training learning' measure found that learning had taken place in all conditions during the training blocks (RTs: ts>9.7; accuracy: ts>6.0).

'Transfer effect'. The ANOVAs on the 'transfer effect' for RTs and the proportion of errors found that the decrement in performance between the end of training and the start of transfer was much larger when novel categories were introduced at transfer (RT difference = 180 ms; accuracy difference = 7.5%) relative to when the categories were repeated (RT difference = 34 ms; accuracy difference = 2.4%; main effect of Category for RTs: F=28.61; BF<0.001 ± 2.7%; main effect of Category for accuracy: F=14.98, BF=0.002 ± 4.5%). The main effect of Stimulus did not reach significance for RTs (F=2.33, BF=2.976 ± 2.7%) but did for accuracy (F=5.27) suggesting that the 'transfer effect' for accuracy was larger for novel stimuli (difference = 6.3%) relative to those stimuli that were repeated from training (difference = 3.5%). However the Bayesian analysis found only anecdotal evidence that removal of the main effect of Stimulus from the model would impair its fit (BF=0.427 ± 4.3%). Nonetheless, stimulus repetition did not influence the 'transfer effect' much when novel categories were introduced at transfer (DsDcSr difference for RTs = 170 ms, accuracy difference = 7.5%; SsDcSr difference for RTs = 189 ms, accuracy difference = 7.4%), but it did have an effect when the categories were repeated from training (DsScSr difference for RTs = 69 ms, accuracy difference = 5.2%; SsScSr difference for RTs = -1 ms, accuracy difference = -0.4%). This pattern of results was confirmed by the significant Stimulus by Category interaction for both RTs (F=6.73) and the proportion of errors (F=5.62), although the Bayesian analyses both found only anecdotal evidence that removal of the interaction from the model would impair its fit (RT: BF=0.550 ± 2.4%; accuracy: BF=0.384 ± 4.4%). Follow-up paired samples t-tests found that the 'transfer effect' was significantly smaller in the DsScSr condition relative to the SsDcSr condition for RTs ($t(39)=3.64$, $p<0.001$, g_{av}=.923, BF=37.68) but not the proportion of errors (t=1.24, BF=0.348) indicating that C-R associations transferred to novel stimuli much easier than S-R associations transferred across classifications.

Taken together, analysis of the 'transfer effect' measure provided strong evidence that learned C-R associations rapidly transfer to novel stimuli whereas the evidence for transfer of learned S-R associations across classifications was much weaker. Nonetheless, that the Stimulus by Category interaction was reliable for both RTs and errors (though the Bayesian analysis only provided anecdotal support for the interaction) suggests that both C-R and S-R learning had contributed to performance in the current experiment.

'Transfer learning'. The ANOVAs on the 'transfer learning' measure found that the improvement in performance through the transfer blocks was much larger when novel categories were introduced at transfer (RT improvement: 169 ms; accuracy improvement: 5.0%) relative to when the categories were repeated (RT improvement: 36 ms; accuracy improvement: 2.0%; main effect of Category for RTs: F=31.46, BF<0.001± 7.1%; main effect of Category for accuracy: F=5.23, BF=0.178 ± 4.2%). Neither the main effect of Stimulus nor the Stimulus by Category interaction reached significance for RTs (Fs<2.2, BFs>2.4) or the proportion of errors (Fs<3.9, BFs>0.61) indicating that, although learning through the transfer blocks was strongly modulated by whether the category templates were novel or repeated from training, transfer learning was largely unaffected by the introduction of novel stimuli.

Interim summary

Close inspection of the results from Experiments 1-3 indicates that, despite some minor statistical differences, the pattern of results is relatively consistent. In the (critical) 'transfer effect' measure, the main effect of Category was reliable in all experiments, but in Experiment 2 the Stimulus by Category interaction (RTs only) also reached significance and in Experiment 3 both the main effect of Stimulus (proportion of errors only) and the Stimulus by Category interaction (RTs and errors) reached significance. However, the Bayesian analyses provided only weak evidence for these effects. Thus, across all experiments, we found strong evidence that learned C-R associations readily transferred to novel stimuli from the same category, whereas the evidence for transfer of learned S-R associations across classifications was much less consistent. In order to determine which effects would withstand more rigorous scrutiny we merged the data from Experiments 1-3 (only the data from the 4 exemplar stimuli conditions in Experiment 2 were included) and re-ran the analyses on the 'transfer effect'. The results from the omnibus ANOVAs are reported in Table 4.

The 'transfer effect' was much larger when novel categories were introduced at transfer (RT = 131 ms, errors = 6.8%) relative to when the categories were repeated from training (RT difference = 17 ms, main effect of Category: F=47.01, BF<0.001 ± 2.7%; accuracy difference = 1.8%, main effect of Category: F=28.25, BF<0.001 ± 2.5%). The 'transfer effect' was also slightly larger for the novel stimuli introduced at transfer (RT difference = 87 ms, accuracy difference = 4.6%) relative to those stimuli that were repeated from training (RT difference = 61 ms, accuracy difference = 4.0%). However, the main effect of Stimulus only reached significance in the RTs (F=6.70) and not the proportion of errors (F<1), and the Bayesian analyses indicated that removal of the main effect of Stimulus from the model would not materially impair its fit (RT: *BF*=1.622 ± 3.1%; errors: *BF*=8.327 ± 3.0%). Taken together these results indicate that whether the categories used at transfer were novel or repeated from training had a much larger impact on performance than whether the stimuli were novel or repeated from training.

When novel categories were introduced at transfer, the 'transfer effect' was slightly smaller for altogether novel stimuli (DsDcSr RT difference = 123 ms, accuracy difference = 5.7%) than for those stimuli that had been classified under a different category during training (SsDcSr RT difference = 139 ms, accuracy difference = 7.8%); by contrast, when the categories were repeated from training, the 'transfer effect' was *larger* for novel stimuli from the same category (DsScSr RT difference = 51 ms, accuracy difference = 3.4%) relative to those stimuli that had been classified under the same category during training (SsScSr RT difference = -17 ms, accuracy difference = 0.1%). This pattern of results was confirmed by the reliable Stimulus by Category interaction for both RTs (F=12.71, *BF*=0.063 ± 3.1%) and the proportion of errors (F=13.11, *BF*=0.040 ± 2.6%).

We also performed several paired-samples t-tests in order to compare the 'transfer effect' in each condition more directly. The 'transfer effect' in the SsDcSr condition was not significantly larger than baseline (DsDcSr) for either RTs (t<1, *BF*=0.163) or errors (t=1.73, *BF*=0.429), indicating that the numerical interference effect due to classifying a familiar stimulus under a novel category was not reliable. Conversely, the 'transfer effect' in the DsScSr condition was significantly smaller than baseline for both RTs (t=3.62, p<0.001, g_{av}=0.462, *BF*=45.04) and the proportion of errors (t=2.04, p=0.044, g_{av}=0.256, *BF*=0.748), indicating that learned C-R associations readily transferred to novel stimuli from the same category. However, the 'transfer effect' in the DsScSr condition was also significantly larger than in the SsScSr condition for both RTs (t=4.61, p<0.001, g_{av}=0.492, *BF*=1430.16) and errors (t=3.51, p<0.001, g_{av}=0.409, *BF*=31.06) indicating that C-R transfer was less than perfect (i.e., when the categories were repeated between training and transfer, novel stimuli were more difficult to classify than familiar stimuli). Critically, the 'transfer effect' was much larger in the SsDcSr condition relative to the DsScSr condition for both RTs (t=4.54, p<0.001, g_{av}=0.538, BF=1109.44) and errors (t=3.36, p=0.001, g_{av}=0.457, BF=19.80) indicating that C-R transfer to novel stimuli was much stronger than S-R transfer across classifications.

In summary, the data from Experiments 1-3 indicate that whether the transfer categories were novel or repeated from training had a much larger effect on performance than whether the transfer stimuli were novel or repeated from training. Nevertheless, when the categories were repeated between training and transfer, transfer of learned C-R associations to novel stimuli was less than perfect. This indicates that C-R bindings alone cannot explain all learning effects. Critically, transfer of learned C-R associations to novel stimuli from the same category was much stronger than transfer of learned S-R associations across classifications.

Experiment 4

The purpose of Experiments 1-3 was to directly compare C-R learning independent of the stimulus and S-R learning independent of the classification. In all experiments we found strong evidence that learned C-R associations readily transfer to novel stimuli whereas the evidence for transfer of learned S-R associations across classifications was much weaker (learning in the SsDcSr condition was even numerically worse than baseline suggesting some carryover of S-C associations from training). That C-R transfer was less than perfect in Experiments 2 and 3 also suggests a possible role of S-C learning in the SsScSr condition where stimulus-category-response bindings were consistent between training and transfer (transfer performance was better when the S-C bindings were consistent between training and transfer relative to when they were inconsistent).

Despite S-C associations receiving considerable attention in the cognitive control and associative learning literature (e.g., Dreisbach, 2012; Horner & Henson, 2009, 2011; Mayr &

Bryck, 2005; Moutsopoulou et al., 2012, 2013, 2015), no researchers have yet directly compared the relative contribution of C-R, S-R and S-C associations to learning. By repeating the relevant responses between training and transfer in Experiments 1-3 it was not possible to take a direct measurement of S-C learning (independent of the response). Therefore, in Experiment 4, we also manipulated whether the relevant responses were repeated between training and transfer. This resulted in a 2 (Stimulus: same vs. different) by 2 (Category: same vs. different) by 2 (Response: same vs. different) design, which allowed us to directly compare the extent to which the various associations outlined above contribute to learning in the current paradigm.

If S-C associations make a valuable contribution to learning in the current paradigm then the benefits of repeating both the stimulus and its classification between training and transfer should expedite learning even if novel responses are introduced at transfer (i.e., the ‘transfer effect’ should be relatively small in the SsScDr condition by comparison to baseline). Experiment 4 therefore allows a direct comparison of learning across the three critical conditions that measure C-R learning independent of the stimulus (DsScSr), S-R learning independent of the classification (SsDcSr) and S-C learning independent of the response (SsScDr).

Method

40 different students from the University of Exeter (32 female) with a mean age of 19.5 years (SD = 2.6) participated for the same reimbursement as Experiments 1-3 (£7 or partial course credits). The target sample size and exclusion criteria were identical to Experiments 1-3, and written informed consent was obtained after the nature and possible consequences of the study was explained.

The apparatus, stimuli, responses and procedure were identical to Experiment 3 (see Figure 2) apart from one critical change: in addition to manipulating whether the category templates were novel or repeated between training and transfer, the relevant response codes used during the transfer blocks could also be novel or repeated from the preceding training block. This resulted in eight conditions in total (DsDcDr, SsDcDr, DsScDr, SsScDr, DsDcSr, SsDcSr, DsScSr, SsScSr). Several other minor changes were made in order to optimize the design of the experiment. As in Experiment 3, the training blocks consisted of 80 trials (four stimuli from each category were presented ten times each), but the transfer blocks consisted of 48 trials (four stimuli from each category were presented six times each). Two training-transfer pairs (consisting of 128 trials each) used novel categories and novel responses at transfer (DsDcDr and SsDcDr conditions), two pairs repeated the categories but used novel responses at transfer (DsScDr and SsScDr conditions), two pairs used novel categories but repeated the responses at transfer (DsDcSr and SsDcSr conditions), and two pairs repeated both the categories and the responses at transfer (DsScSr and SsScSr conditions). This resulted in a maximum of 80 observations per condition during training blocks and 48 observations per condition during transfer blocks (total number of experimental trials = 1024). As in Experiments 1-3, whether the transfer block used the same or different categories to those used during training alternated through the experiment (e.g., same categories, different categories, same categories, different categories...). Whether the transfer block used the same or different responses to those used during training changed on every second training-transfer pair (e.g., same responses, same response, different responses, different responses...). The order of conditions was Latin square balanced over participants.

There are 24 unique 4-digit response codes that start and end with 5 and where each intervening digit is on adjacent keys on the numeric keypad of a standard keyboard. All 24 response codes were used during the experimental session so it was necessary to use a different format for the practice session to avoid any unwanted transfer effects. The first practice block (categorization practice; 80 trials) was identical to a training block from the experimental session, apart from two differences: (1) participants were required to identify category membership by pressing either the 'a' or 'l' key on a standard keyboard with their left or right index finger respectively; (2) the response was not displayed on screen as it was entered, but immediate feedback was presented for 1000 ms as soon as a response was made. Each trial in the second practice block (response practice; 48 trials), started with a blank screen (500 ms) followed by a 4-digit response code presented centrally in black Ariel font size 30. Every possible code was presented twice through the block (order pseudo-radomized). The participant was then required to enter the code as quickly as possible while minimizing errors using their preferred index finger and the numeric keypad on a standard keyboard (all participants used their right hand). The response code appeared on the screen as it was entered and immediate feedback was given in the form of the correct response code presented in either green (correct) or red (error) for 1000 ms. Participants were informed that the practice blocks were distinct from the experimental session that followed. Each practice block was preceded by written instructions regarding their procedure presented on the screen and was initiated by pressing the space bar once the instructions had been read. The data from the practice session were not analyzed.

As in Experiments 1-3, trials with RT (first digit) <100 ms (0.06%) and trials with RT (last digit) >5000 ms (0.35%) were omitted from all analyses, and error trials were omitted from RT analyses. It was necessary to replace the data from four participants because they had <50% of the maximum possible observations in at least one condition (excluding the practice blocks) following the above data cleaning procedures.

For each score and dependent variable (RT and error rate), we performed a separate ANOVA with the factors Stimulus (same, different), Category (same, different), and Response (same, different). We also report Bayes factors and effect sizes (generalized eta squared) for all relevant effects/interactions (calculated in the same way as for Experiments 1-3). As in Experiments

1-3, one-sample t-tests were performed on the 'training learning' measure to confirm that learning had taken place in each condition during the training blocks. Several additional paired-samples t-tests were performed to directly compare the relative contribution of C-R, S-R and S-C associations to learning. Bayes factors and effect sizes (g_{av}) for these comparisons were calculated in the same way as for Experiments 2 and 3.

Results and Discussion

Mean RTs and the proportion of errors made in the training and transfer blocks (separately for conditions where the responses used at transfer were novel or repeated from training) are plotted as a function of condition (DsDcDr, DsScDr, SsDcDr, SsScDr, DsDcSr, DsScSr, SsDcSr, SsScSr) and stimulus repetition in Figure 8. The mean 'transfer effect' for RTs and the proportion of errors are plotted as a function of condition in Figure 5. The results from the omnibus ANOVAs are reported in Table 5.

'Training learning'. The ANOVA on the 'training learning' measure for RTs found a reliable three-way interaction (F=6.05). However, the Bayesian analysis found only anecdotal evidence that removal of the interaction from the model would impair its fit (BF=0.468 ± 13.2%). None of the other effects or interactions approached significance (Fs<2.9, BFs>2.1), neither did any of the effects or interaction in the ANOVA on the proportion of errors (Fs<1, BFs>3.3). Taken together, these results indicate that there was no systematic modulation of the amount learned during training across the conditions. The one-sample t-tests on the 'training learning' measure found that learning had taken place in all conditions during the training blocks (RTs: ts>5.3; accuracy: ts>5.3).

'Transfer effect'.

Response times. The ANOVA on the 'transfer effect' for RTs found that the decrement in performance between the end of training and the start of transfer was much larger when novel categories were introduced at transfer (difference = 148 ms) relative to when the categories were repeated from training (difference = 50 ms; main effect of Category: F=10.33, BF<0.001 ± 4.7%). The 'transfer effect' was also larger when novel responses were introduced at transfer (difference = 132 ms) relative to when they were repeated from training (difference = 66 ms; main effect of Response: F=10.23, BF=0.025 ± 4.9%). The main effect of Stimulus did not approach significance (F=2.02, BF=7.90 ± 26.0%). Taken together these results indicate that transfer performance was strongly modulated by whether the categories and/or the responses used at transfer were novel or repeated from training, whereas the introduction of novel stimuli at transfer had little effect.

When the categories changed at transfer, the 'transfer effect' was numerically smaller for altogether novel stimuli (DsDcSr, DsDcDr average difference = 133 ms) than for those stimuli that had previously been classified according to a different category (SsDcSr, SsDcDr average difference = 163 ms). On the other hand, when the category templates were repeated between training and transfer, the transfer effect was *larger* for novel stimuli (DsScSr, DsScDr average difference = 83 ms) than for those stimuli that had been previously classified under the same category (SsScSr, SsScDr average difference = 18 ms). This pattern of results was supported by a significant Stimulus by Category interaction (F=10.62, BF=0.338 ± 5.6%). A follow-up paired samples t-test found that the 'transfer effect' was significantly smaller novel stimuli classified under a familiar category (i.e., the DsScSr and DsScDr conditions) relative to familiar stimuli classified under a novel category (i.e., the SsDcSr and SsDcDr conditions; t=2.33, p=.025, g_{av}=.533, BF=1.88) suggesting a bias toward transfer of learned C-R associations to novel stimuli over the transfer of learned S-R associations across classifications.

The omnibus ANOVA also found a reliable interaction between Stimulus and Response (F=4.26), but the Bayesian analysis found substantial evidence that removing the interaction from the model would *not* impair its fit (BF=3.217 ± 6.2%). Neither the Category by Response interaction (F=2.36, BF=1.604 ± 8.4%) nor the three-way interaction (F<1, BF= 4.367± 6.1%) approached significance.

In order to determine the relative contribution of C-R, S-R and S-C associations to learning, the 'transfer effect' in each critical condition (DsScSr difference = 41 ms; SsDcSr difference = 131 ms; SsScDr difference = 75 ms) was compared to baseline (DsDcDr difference = 135 ms). The 'transfer effect' was reliably smaller than baseline in the DsScSr condition ($t(39)$=2.36, p=.023, gav=.552, BF=2.0), but not in either the SsDcSr condition (t=0.07, BF=0.17) or the SsScDr condition (t=1.59, BF=0.54) indicating that learned C-R associations readily transferred to novel stimuli whereas the evidence for transfer of learned S-R associations to novel classifications and S-C associations to novel responses was much weaker.

Accuracy. The ANOVA on the 'transfer effect' for the proportion of errors also found that the decrement in performance between the end of training and the start of transfer was larger when novel category templates were introduced (difference = 10.0%) relative to when the category templates were repeated (difference = 3.2%; main effect of Category: F=42.98, BF<0.001 ± 7.2%). The 'transfer effect' was also larger for novel stimuli (difference = 8.0%) relative to those stimuli that were repeated from training (difference = 5.3%; main effect of Stimulus: F=5.89). However, the Bayesian analysis found only anecdotal evidence that removal of the main effect of Stimulus from the model would impair its fit (BF=0.763 ± 5.7%). The main effect of Response did not approach significance (F=1.42, BF=3.136 ± 6.7%). Taken together these results provide further support for the notion that learning in the current paradigm is modulated by whether the category templates used at transfer were novel or repeated from training, whereas the introduction of novel stimuli or responses had a smaller effect on performance.

As in the RTs, the ‘transfer effect’ was numerically smaller for altogether novel stimuli (DsDcSr, DsDcDr average difference = 9.9%) than for those stimuli that had previously been classified according to a different category (SsDcSr, SsDcDr average difference = 10.1%). On the other hand, when the categories were repeated between training and transfer, the ‘transfer effect’ was *larger* for novel stimuli (DsScSr, DsScDr average difference = 6.0%) than for those stimuli that had previously been classified (SsScSr, SsScDr average difference = 0.4%). This pattern of results was supported by a significant Stimulus by Category interaction (F=6.44), although the Bayesian analysis found only anecdotal evidence that removal of the interaction from the model would impair its fit (BF=0.510 ± 12.7%). A follow-up paired samples t-test found that the ‘transfer effect’ was significantly smaller for novel stimuli classified under a familiar category (i.e., the DsScSr and DsScDr conditions) relative to familiar stimuli classified under a novel category (i.e., the SsDcSr and SsDcDr conditions; $t(39)=2.32$, $p=.026$, $g_{av}=.508$, BF=1.84) further supporting the notion of a bias toward transfer of learned C-R associations to novel stimuli over the transfer of learned S-R associations across classifications. None of the other interactions approached significance (Fs<1, BFs>4.4).

As with the RTs, several additional paired-samples t-tests were performed to determine the relative contribution of C-R, S-R and S-C associations to learning. The ‘transfer effect’ in each critical condition (DsScSr difference = 4.0%; SsDcSr difference = 10.2%; SsScDr difference = 1.3%) was compared to the baseline condition (DsDcDr difference = 10.7%) to determine whether transfer performance was materially improved by each kind of learned association. The ‘transfer effect’ was reliably smaller than baseline in the DsScSr condition ($t(39)=2.76$, $p=.009$, $gav=.639$, BF=4.61) and the SsScDr condition ($t(39)=4.06$, $p<.001$, $gav=.922$, BF=115.20), but not in the

SsDcSr condition (t=0.17, BF=0.17) indicating that learned C-R associations readily transferred to novel stimuli and learned S-C associations readily transferred to novel responses whereas the evidence for transfer of learned S-R associations to novel classifications was much weaker.

'Transfer Learning'. The ANOVAs on the 'transfer learning' measure found that the improvement in performance through the transfer block was much larger when novel categories were introduced at transfer (RT improvement = 131 ms, accuracy improvement = 6.8%) relative to when the categories were repeated (RT improvement = 68 ms, accuracy improvement = 3.3%; main effect of Category for RTs: F=10.46, BF=0.001 ± 13.3%; main effect of Category for accuracy: F=15.97. BF=0.029 ± 9.5%). The 'transfer learning' measure was also larger when novel responses were introduced at transfer (RT improvement = 139 ms, accuracy improvement = 6.6%) relative to when the responses were repeated from training (RT improvement = 60 ms, accuracy improvement = 3.5%; main effect of Response for RTs: F=19.68, BF<0.001 ± 10.2%; main effect of Response for accuracy: F=8.28, BF=0.098 ± 9.6%). The main effect of Stimulus did not approach significance for either RTs (F<1, BF=10.166 ± 40.1%) or errors (F<1, BF=8.619 ± 7.7%).

The Stimulus by Category interaction was reliable for the proportion of errors (F=7.62, BF=0.190 ± 10.4%) mirroring the pattern of results in the 'transfer effect' measure; however, the interaction did not approach significance for RTs (F<1, BF>4.2). None of the other interactions were reliable for RTs (Fs<1.1, BFs>4.0) or the proportion of errors (Fs<1, BFs>4.1) indicating that transfer learning was modulated by whether the categories and/or responses were novel or repeated from training, but not by the introduction of novel stimuli.

Summary. Experiment 4 was designed to directly compare the relative contribution of C-R, S-R and S-C associations to learning in the current paradigm. There was clear evidence that learned C-R associations transferred to novel stimuli from the same category, whereas the evidence for transfer of learned S-R associations across classifications was much weaker and the evidence for transfer of learned S-C associations to novel responses was less consistent (the 'transfer effect' was reliably smaller than baseline in the proportion of errors but not in the RTs). That S-C transfer was evident (particularly in the accuracy data) suggests the importance of S-C learning in the current paradigm; however this effect is outweighed by the importance of C-R learning (which was evident in both dependent variables). Although it is possible that this pattern of results is dictated by some critical features of the current paradigm, that we report a situation where C-R associations make a substantial contribution to learning (above the contributions of both S-R and S-C associations) has potential implications far beyond the scope of this article (we return to this point in the General Discussion).

Experiment 5

As discussed in the Introduction, psychologists often instruct participants in the C-R bindings (e.g., odd number → left hand response, even number → right hand response; living → index finger, man-made → middle finger). Likewise 'real life' instructions also often focus on categories (e.g., edible fruit → eat, rotten fruit → discard; domesticated animal → approach, wild animal → avoid). Experiments 1-4 were designed to determine whether under such circumstances, the learning of C-R associations (independent of the specific stimulus) contribute to performance. Therefore, we mentioned the category structure in these experiments (i.e., participants were told that on each trial they would be presented with a dot-pattern and that they should decide which of the two relevant category templates the pattern most closely resembled before entering the correct code for the category) and the category templates were presented prior to each experimental block. Furthermore, immediate feedback consisted of the category template (Experiments 1 and 2) or the category name (Experiments 3 and 4) presented alongside the correct response code. Conversely, the stimuli themselves were not presented during the instructions, were overlaid by the (much more salient) category template during immediate feedback in Experiments 1 and 2, and were not presented at all during immediate feedback in Experiments 3 and 4. Our results indicate that participants adopted a strategy that utilized the category structures imposed on them and were able to rapidly transfer learned C-R associations to novel stimuli from the same category, but found it relatively difficult to transfer learned S-R associations across classifications.

Evidence of C-R learning under any conditions is noteworthy, simply because of its contribution to the growing bank of evidence that learning need not be strictly limited to S-R associations. Nonetheless, it seems important to determine whether the pattern of results reported in Experiments 1-4 are limited to situations where the instructions describe the C-R bindings and not the S-R bindings.

To this end, in Experiment 5 we generated stimuli based on the category templates, but we omitted any reference to the categories in the general instructions, presented the stimuli (and not the category templates/names) alongside the correct response codes during the pre-block instructions, and immediate feedback consisted of only the correct response code presented in either green (correct) or red (error). The stimuli, category templates used to generate the stimuli, and the design of the experiment were the same as those used in Experiment 3 but we never showed the category templates to the participants. Thus, in Experiment 5, the only difference between conditions in which the categories were the same or different to the categories used during training was whether the novel stimuli introduced at transfer were perceptually similar to those observed during training (i.e., they were based on the same category template) or were perceptually distinct (i.e., they were based on a different category template). In other words, both the SsDcSr condition and the SsScSr condition used identical stimuli at transfer, but the former were accompanied by novel stimuli based on novel category templates (i.e., the DsDcSr condition) whereas the latter were accompanied by novel stimuli based on the same category templates used to draw the stimuli that they replaced (i.e., the DsScSr condition). As can be seen in Figure 1, this implied that the overall similarity between training and transfer stimuli was higher for the SsScSr condition than the SsDcSr condition.

If our observation that C-R associations are an important contributor to the early stages of learning can largely be explained by the instructions (and the features of the paradigm) which encouraged participants to associate responses to categories rather than stimuli, then changing the instructions (and eliminating those features) should also eliminate the relative benefits of C-R learning (i.e., transfer of C-R associations to novel stimuli). On the other hand, if changing the instructions does not eliminate the bias toward C-R learning then we can conclude that categories are an important contributor to learning, even when they are not overtly instructed or mentioned[9].

Method

40 different students from the University of Exeter (34 female) with a mean age of 20.4 years (SD = 4.3) participated for the same reimbursement as Experiments 1-4 (£7 or partial course credits). The target sample size and exclusion criteria were identical to Experiments 1-4, and written informed consent was obtained after the nature and possible consequences of the studies were explained.

The apparatus, stimuli, responses and procedure were identical to Experiment 3 apart from two critical changes. First, the instructions screen at the start of each block displayed all eight stimuli in two rows of four stimuli each. The four stimuli that were derived from the same category template and that shared the same response were presented on either the left or right of the screen with the correct response code presented directly below each stimulus (see Figure 2). The relative location of each stimulus within the group was pseudo-randomly selected for each block. Second,

[9] It should be noted that the categorization literature shows that participants will learn categories even when the templates are not shown (e.g., Wills et al., 2006), though this process can take a relatively long time by comparison to the rapid acquisition of categories in the current experiments. Furthermore, it could be argued that in many categorization studies participants are explicitly instructed to categorize stimuli which might encourage such a strategy from the outset.

immediate feedback on each trial consisted of only the correct response code presented centrally in either green (correct) or red (error). The general instructions at the start of the experimental session were also edited to reflect these changes to the procedure.

As in Experiments 1-4, trials with RT (first digit) <100 ms (0.20%) and trials with RT (last digit) >5000 ms (0.53%) were omitted from all analyses, and error trials were omitted from RT analyses. It was not necessary to replace the data from any participants due to poor performance. The analyses were identical to those performed in Experiment 1.

Results and Discussion

Mean RTs and the proportion of errors made in the training and transfer blocks are plotted as a function of condition (DsDcSr, DsScSr, SsDcSr, SsScSr) and stimulus repetition in Figure 9. The mean ‘transfer effect’ for RTs and the proportion of errors are plotted as a function of condition in Figure 5. The results from the omnibus ANOVAs are reported in Table 6.

‘Training learning’. None of the effects or interactions in the ANOVAs on the ‘training learning’ measure was reliable for RTs (*F*s<1, *BF*s>3.6) or the proportion of errors (*F*s<3.6, *BF*s>1.0) indicating that the improvement in performance through training was comparable in all conditions. The one-sample t-tests on the ‘training learning’ measure found that learning had taken place in all conditions during the training blocks (RTs: *t*s>7.7; accuracy: *t*s>5.8).

‘Transfer effect’. The ANOVAs on the ‘transfer effect’ for RTs and the proportion of errors found that the decrement in performance between the end of training and the start of transfer was much larger for novel stimuli that were introduced at transfer (RT difference = 97 ms; accuracy difference = 5.6%) relative to those stimuli that were repeated from training (RT difference = 45 ms, accuracy difference = 0.3%; main effect of Stimulus for RTs: F=14.44, BF=0.060 ± 2.7%; main effect of Stimulus for accuracy: F=20.10, BF=0.006 ± 3.2%). Neither the main effect of

Category nor the Stimulus by Category interaction reached significance for either RTs (*F*s<2.8, *BF*s>0.7) or the proportion of errors (*F*s<3.6, *BF*s>0.9) suggesting that learned S-R associations readily transferred to a novel context whereas the evidence for C-R learning was sparse.

'Transfer learning'. The ANOVAs on the 'transfer learning' measure found that the improvement in performance through the transfer blocks was larger for novel stimuli (RT improvement = 89 ms; accuracy improvement = 6.3%) relative to the stimuli that were repeated from training (RT improvement = 58 ms; accuracy improvement = 1.9%). The difference was reliable in the response accuracy data (main effect of Stimulus: *F*=19.65, *BF*=0.003 ± 3.1%), but not in the RTs (main effect of Stimulus: *F*=3.76, *BF*=1.163 ± 16.9%). Neither the main effect of Category nor the Stimulus by Category interaction approached significance for RTs (*F*s<1.6, *BF*s>1.6) or the proportion of errors (*F*s<2.5, *BF*s>1.6) indicating that, although learning through the transfer blocks was strongly modulated by whether the stimuli were novel or repeated from training, transfer learning was largely unaffected by the introduction of stimuli based on novel category templates.

Summary. The results from Experiment 5 indicate the importance of instructions to learning. When the instructions emphasized the C-R bindings (as in Experiments 1-4) the evidence for transfer of learned C-R associations to novel stimuli outweighed the evidence for transfer of learned S-R associations across classifications, but when the instructions emphasized the S-R bindings the evidence for transfer of learned S-R associations to a novel context outweighed the evidence for C-R learning. It is doubtful that the participants were able to memorize all eight S-R bindings during the 15-second pre-block instructions phase. However, displaying all eight stimuli and the correct response associated with each at least provided participants with the opportunity to adopt one of two strategies: they might have attempted to memorize a subset of the S-R bindings and then used the immediate feedback to learn the rest; alternatively they might have attempted to look for commonalities between the stimuli that shared a response (i.e., attempted to infer the categories). The absence of a main effect of category repetition in the transfer blocks, is inconsistent with the second explanation. Thus, we propose that participants primarily relied on an 'exemplar' strategy. That response accuracy was better than chance from the outset suggests that the participants were successful in memorizing at least some of the S-R bindings presented during the pre-block instructions phase.

Intriguingly, the 'transfer effect' was numerically larger in the SsDcSr condition (RT difference = 66 ms; accuracy difference = 3.1%) relative to the SsScSr condition (RT difference = 51 ms; accuracy difference = 1.7%). Paired-samples t-tests found that the difference did reach significance for the proportion of errors ($t(39)=2.19$, $p=0.034$, $g_{av}=0.557$, $BF=1.452$), but not RTs ($t=1.60$, $BF=0.552$). We can only speculate that this surprising result could be due to the relative similarity between the novel stimuli introduced at transfer and the stimuli that they replaced – when the novel stimuli were based on a familiar category template, classification of the stimuli that were repeated from training was slightly improved relative to when the novel stimuli were based on a novel category template. That the Bayesian analyses found only anecdotal evidence in either direction indicates that additional research is necessary to further investigate this claim.

Finally, it could be argued that the C-R bias observed in Experiments 1-4 is, at least in part, due to category (and response) novelty being manipulated between blocks (i.e., were consistent within a block) whereas stimulus novelty was manipulated within a block (some transfer stimuli were novel and some were repeated from training). However, this relationship was also true in Experiment 5, where the C-R bias was not found (indeed, evidence of C-R transfer at all was sparse). If the C-R bias reported in Experiments 1-4 was due to some procedural manipulation inherent in the design of the experiments then the same pattern of results should have been found in

Experiment 5, where the design was identical to Experiment 3 apart from the format of the pre-block instructions.

General discussion

The main aims of the current study were twofold: (1) we wanted to investigate the relative contribution of C-R associations (independent of the stimulus), S-R associations (independent of the classification) and S-C associations (independent of the response) to performance in an instructed category-learning paradigm; and (2) we wanted to determine the extent to which instructions can modulate the relative benefits of each type of association. Our findings are largely consistent with the recent interest in associations other than simple S-R learning within the cognitive control and associative learning literature (e.g., Hazeltine & Schumacher, 2016; Henson et al., 2014) and make a unique contribution by directly comparing the relative importance of C-R, S-R and S-C associations in categorization studies that use artificial stimuli/categories and complex responses. Furthermore, our findings contribute to the growing body of research concerned with instructed learning by highlighting the importance of framing instructions in such a way as to encourage learning (and transfer) of the desired content (associations).

In Experiment 1, we found strong evidence for C-R learning and transfer, but only weak evidence for S-R learning (independent of the classification). In Experiment 2, we contrasted a 4 exemplar stimuli condition with a 16 exemplar stimuli condition. There was equal opportunity to learn the C-R associations in both conditions, but if S-R associations were important in the current design (when the instructions emphasized the C-R bindings) then dramatically increasing the number of S-R bindings to learn should have further extended the bias toward C-R learning. The absence of a meaningful difference between these two conditions in Experiment 2 suggests that C-R associations are central to learning in the current paradigm and are independent of the number of exemplar stimuli per category. Experiment 3 ruled out the possibility that incidental associations were formed between the (perceptual) category template and the correct response presented during immediate feedback (a potential confound in Experiments 1 and 2). The relative contribution of C-R, S-R and S-C associations to learning was directly examined in Experiment 4 by also manipulating whether the responses used at transfer were novel or repeated from training. Although there was some evidence that learned S-C associations transferred to novel responses (the 'transfer effect' was smaller in the SsScDr condition relative to baseline for response accuracy, but not RTs), it was outweighed by the evidence for transfer of learned C-R associations to novel stimuli (the 'transfer effect' was reliably smaller in the DsScSr condition relative to baseline for both dependent variables). The current experiments therefore indicate the particular importance of C-R associations in learning and transfer and indicate that, at least when the instructions mention the categories, S-R associations rely on a consistent mapping between the stimulus, the category and the response (see also Moutsopoulou et al., 2012, 2013, 2015).

Combined, the results from Experiments 1-4 provided strong evidence that, when the category templates are presented at the start of each experimental block, transfer of C-R associations to novel stimuli was much stronger than transfer of S-R associations across classifications or S-C associations to novel responses. However, C-R transfer was also found to be less than perfect (the 'transfer effect' was larger in the DsScSr condition than in the SsScSr condition) suggesting that consistent S-C bindings also expedite learning. Furthermore, that we observed a small (albeit unreliable) numerical performance cost in the 'transfer effect' measure for the SsDcSr condition relative to baseline in Experiments 2-4 is suggestive of interference when learned S-R associations are performed under a novel classification (though further research is necessary to confirm this). This could, at least in part, be due to carryover of the S-C associations formed during training, but Experiment 4 found that S-C transfer did not modulate performance to the same extent as transfer of C-R associations. This observation is also consistent with Mayr and Bryck's (2005) finding that the performance benefits associated with learned S-R bindings are category (rule) specific in a task switching context (see also Moutopoulou et al, 2012, 2013, 2015; though, for an example of S-R effects independent of the category/rule in a go/no-go paradigm see Verbruggen & Logan, 2008).

Experiment 5 provided evidence that the C-R learning bias observed in Experiments 1-4 was at least in part due to the framing of the instructions – a simple modulation of the instructions was enough to shift that bias towards S-R learning. The surprising observation that the 'transfer effect' was numerically larger in the SsDcSr condition relative to the SsScSr condition was likely due to the perceptual similarity between the training and test stimuli in these conditions (as discussed above). However, it should be noted that previous work indicates that participants might learn categories even without instructions. For example, categorization experiments commonly require the participants to infer the categories without being presented with the templates (e.g., Wills et al., 2006) and sometimes without any feedback (e.g., Wills & McLaren, 1998), though it could be argued that instructing participants to categorize the stimuli is enough to direct their attention toward the category structures/rules used to draw the stimuli. Also, Collins and Frank (2013) suggest that participants spontaneously build task-set structure into learning problems even when not instructed to do so (see also Dreisbach, 2012). Presumably, searching for coherent and useful category structures and/or rules without any instruction will take considerably longer than the relatively short time participants took to learn the categories when presented with the templates as in the current Experiments 1-4. Further research is necessary to clarify whether inferring the underlying category structures used to produce groups of stimuli that require the same response is the norm and how long it takes to develop useful category structures in this manner that can direct performance.

Where other researchers interested in cognitive control, associative learning and/or automaticity have investigated C-R associations, they have tended to use well-learned taxonomic categories and/or extensive training regimes. The current findings extend this research to include artificial (i.e., altogether novel) categories rapidly acquired over a relatively short sequence of stimulus/category repetitions via instruction. This extension is critical in developing our understanding of how cognitive control (and metacognition, cf. Chein & Schneider, 2012) direct the early stages of instructed learning. Instructed rule-based learning relies on C-R associations that elicit reasonably accurate performance on the first trial in a run and which rapidly develop over the course of a few subsequent trials/instances. These associations easily transfer to novel stimuli thereby expediting subsequent learning. It is likely that extensive training with a limited set of stimuli would eventually speed specific S-R associations (cf. Logan's, 1988, Instance Theory) as well as C-R associations (cf. Collins & Frank, 2013) by encouraging the direct retrieval of the relevant response for the stimulus (or category) from memory. However, our findings would suggest that such stimulus-specific speeding might be limited to S-R associations classified under the same category and such benefits would not necessarily transfer across classifications (see also Mousopoulou et al., 2012, 2013, 2015), whereas the speeding of responses due to learned C-R associations is likely to transfer to novel stimuli from the same category with ease – a view that is mirrored by the general consensus in the transfer literature that abstract knowledge readily transfers to novel tasks that share an underlying structure but which differ in their surface features, whereas an over-emphasis on surface features can sometimes hinder transfer (e.g., Day & Goldstone, 2012).

Much of the research investigating various forms of learning mentioned in the introduction (e.g., Cohen-Kdoshay & Meiran, 2007, 2009; Horner & Henson, 2009, 2011; Liefooghe et al., 2012, 2013; Meiran et al., 2015a, 2015b; Mousopoulou et al., 2012, 2013, 2015) used pictures of everyday items or letters/numbers as well as relatively common classifications and simple responses. It is yet to be seen whether the C-R bias we report here would also be found in experiments with similarly familiar stimuli and/or simple responses. The motivation to use highly artificial dot-pattern stimuli and geometric category templates in the current experiments was twofold: first we intended to provide conditions under which C-R learning could be beneficial because the stimuli were relatively difficult to distinguish, but the categories were relatively easy to distinguish; second, we wanted to investigate the very early stages of learning without the potential interference of prior learning, so we had to create abstract patterns that were less likely to have prior representations/associations. Likewise, complex responses were used in order to minimize any potential interference formed by re-using the same response actions across multiple classifications. Although it is possible that the C-R bias reported in Experiments 1-4 is due to some/all of these features of the paradigm, that we found no such bias in Experiment 5 (which used an identical procedure and stimuli/categories/responses) suggests that the critical feature is the framing of the instructions rather than some feature of the paradigm per se. Although further research is necessary to determine whether the bias toward C-R learning generalizes beyond the artificial categories/stimuli used in the current experiments, some preliminary data collected in our lab suggest that rapid transfer of learned C-R associations to novel stimuli from the same category is also found in paradigms that use familiar stimuli/classifications (see also Cohen-Kdoshay & Meiran, 2007, 2009).

The results from Experiment 5 indicate that the C-R bias is, at least in part, due to the framing of the instructions. The rapid transfer of learned C-R associations was present in all experiments in which the instructions explicitly emphasized the C-R bindings (Experiments 1-4) but when the instructions emphasized the S-R bindings (as in Experiment 5), the bias shifted to S-R learning. S-R instructions are often used in simple two-choice tasks (e.g., Cohen-Kdoshay & Meiran, 2007, 2009; Meiran et al., 2015a, 2015b; Horner & Henson, 2009, 2011; Liefooghe et al., 2012, 2013; Mousopoulou et al., 2012, 2013, 2015), whereas C-R instructions are more regularly used in studies that require larger stimulus sets (e.g., task-switching studies). In everyday life, instructions can also emphasize S-R or C-R mappings. Our study indicates that how these instructions are framed, will influence learning (see also Pereg & Meiran, 2017, who concluded that instructions determined learning in a task-switching paradigm).

It is also yet to be seen whether our observation that participants did not display strong performance benefits when applying learned S-R associations to stimuli classified under a different category in Experiments 1-4 will generalize to conditions with extensive training where there is more potential for specific S-R associations to be strengthened. However, it should be noted that extensive training would also provide more opportunity for C-R associations to form either via instruction (as in the current experiments) or spontaneously (e.g., Collins & Frank, 2013, Dreisbach, 2012). Further research is necessary to clarify these points.

Implications for the study of sequential effects in the control literature

The cognitive control system can adjust or reconfigure lower-level systems between trials, and it is generally assumed that sequential effects in various cognitive control paradigms reflect such control adjustments. Researchers in this domain regularly control for S-R repetitions between trials (to control for learning confounds), but our findings indicate that C-R repetitions should also be considered.

In interference tasks (such as the Stroop or flanker task), the congruency effect (i.e., the difference between incongruent and congruent trials) is usually smaller after an incongruent trial than after a congruent trial (e.g., Gratton, Coles, & Donchin, 1992). This decrease is usually attributed to increased control after stimulus- or response conflict. In task-switching experiments, performance is usually worse when the current task is different to that performed on the previous trial relative to when the task repeats – the 'switch cost' (e.g., Kiesel et al., 2010; Monsell, 2003; Vandierendonck, Liefooghe, & Verbruggen, 2010). The switch cost is often interpreted as indexing active reconfiguration of the attentional or response settings. Finally, in the stop-signal paradigm (which measures response inhibition), response latencies are often longer after a stop-signal trial than after a no-signal trial, which could indicate that subjects strategically alter the balance between going and stopping after a stop-signal trial (e.g. Bissett & Logan, 2011).

However, some researchers have questioned whether such sequential effects are necessarily due to top-down control adjustments. For example, congruency sequence effects (Duthoo, Abrahamse, Braem, Boehler, & Notebaert, 2014; Egner, 2008) and sequential effects in response-inhibition paradigms (e.g. Verbruggen, Logan, Liefooghe, & Vandierendonck, 2008) are partly due to the retrieval of stimulus-specific associations from memory (e.g., stimulus-response or stimulus-outcome associations). Stimulus- or cue repetition effects also play an important role when switching between tasks. For example, Logan and Bundesen (2003) argued that a portion of the 'task' switch cost can be explained by a cue repetition *benefit* (if there is a 1:1 cue:task ratio then task repetitions/switches are perfectly confounded with cue repetitions/switches). Others have indicated the importance of stimulus-task associations (e.g., Allport & Wylie, 2000; Moutsopoulou et al., 2012, 2013, 2015; Waszak et al., 2003; Wylie & Allport, 2000). In an extension of this work, Schmidt and Liefooghe (2016) have recently demonstrated that a large portion of the switch cost can be explained by incidental transition effects (e.g., cue repetitions, stimulus repetitions and response repetitions).

To control for such 'bottom-up' effects, researchers typically exclude stimulus repetitions from their design or include 'stimulus repetition' as a factor in their analyses. However, the present study indicates that this may not be sufficient to separate 'associative learning' effects from 'top-down' control effects. After all, we show that categories can also become associated with responses, and that C-R associations can transfer well to novel stimuli. Consistent with this idea, Logan and colleagues have also advocated the notion that low-level priming effects in the task-switching literature are not necessarily limited to specific instances, but can be generalized to other, conceptually related, instances (e.g., Arrington & Logan, 2004; Arrington, Logan, & Schneider, 2007; Logan & Bundesen, 2003; Logan & Schneider, 2006).

Thus, the current findings suggest that researchers interested in cognitive control experiments should also consider learned C-R associations when interpreting their results. Like Schmidt and Liefooghe (2016), we do not wish to suggest that all processes traditionally labeled 'cognitive control' can be explained by learning of various bindings. But we do urge researchers to acknowledge that a large part of the behavioral effects observed in traditional cognitive control experiments might represent learning of specific associations, including associations between more abstract (cue/stimulus/response category) representations.

The nature of category representation

The debate regarding the nature of category representation has been long and contentious and continues to this day (e.g., Pothos & Wills, 2011). Nonetheless, it is important to note that the current results cannot distinguish between generalization by incorporation of novel stimuli into the category (i.e., the formation of a category-level representation that becomes associated with a response) and generalization due to perceptual similarity between the stimuli (i.e., exemplar models of categorization). It may not be necessary to invoke an association between a category-level representation (which could be a perceptual prototype or a verbal mediator) and the response in order to explain the generalization of learning to novel stimuli from the same category in the current design. Because the novel stimuli introduced during transfer were perceptually similar to other members of the same category and perceptually distinct from members of the other category, generalization from training to transfer could be explained by similarity to other exemplars from the same category. When a novel exemplar is very similar to the other members of the category, old S-R associations might be retrieved, leading to a transfer benefit for the novel category items. Thus, categories become associated with responses via the formation of multiple S-R associations, rather than via the formation of an association between the response and a category prototype/template/mediator.

We cannot distinguish between these options using the current design because exemplar stimuli from a given category are perceptually similar to both the category template and the other exemplars from the category. Although categorization researchers (e.g., Maddox et al., 2010; Wills et al., 2006) continue to attempt to distinguish these two possibilities, further research is necessary to clarify whether C-R generalization in the current experiments is due to the rapid incorporation of novel stimuli into the category structure or simply due to similarity between exemplars.

Whether the rapid C-R transfer we report in conditions where the instructions emphasize the C-R bindings (Experiments 1-4) is due to perceptual similarity between exemplars or the formation of a category-level structure (such as a category prototype or a 'mediator'), we would assume that the C-R bias reported in these experiments was likely the result of presenting the category templates during the pre-block instructions. Given that the category templates provided were highly effective in determining the category membership of each stimulus, the participants were offered precisely the relevant information needed to reduce processing/memory demands by adopting a more general rule-based approach to learning which made it easy for novel stimuli to be integrated into the instructed category structure (whether the stimuli were introduced during training or transfer). The instructions to participants even suggested that they should determine to which category the stimulus belonged and then enter the correct response code for that category. By contrast, without the category templates or instructions suggesting an inherent category structure to the stimuli (as in

Experiment 5), the participants would have been faced with two options: (1) attempt to discern the 'correct' category template (which would have increased processing demands, but might have been worth pursuing if each classification task was performed over a longer duration); or (2) try to learn the S-R bindings (which would require greater memory demands, but may have resulted in better short-term outcomes). Either way, more emphasis would have been placed on the S-R bindings initially because that information was provided in the pre-block instructions whereas no information regarding the categories was provided at all[10].

Conclusion

In conclusion, the current report has highlighted the importance of instruction in learning and the transfer of learned material to a novel context. The experiments reported here provide strong evidence that, as is typical in experimental psychology and in everyday life, when instructions emphasize the C-R bindings, C-R associations can be an important contributor to learning and the benefits of that learning can rapidly transfer to novel stimuli from the same category. Conversely, the evidence for rapid transfer of learned S-R associations across classifications and S-C associations to novel responses was relatively weak. However, when the instructions emphasize the S-R bindings, the bias is reversed resulting in rapid transfer of learned S-R associations to a novel context but scant evidence of C-R learning.

References

Allenmark, F., Moutsopoulou, K., & Waszak, F. (2015). A new look on S-R associations: How S and R link. *Acta Psychologica*, 160, 161-169.

Allen, S. W., & Brooks, L. R. (1991), Specializing the operation of an explicit rule. *Journal of*

[10] Theoretically, it would be possible to infer the categories during the pre-block instructions, and some participants reported that they attempted to do so, but our findings suggest that this was not the norm.

Experimental Psychology: General, 120 (1), 3-19.

Allport, D. A., & Wylie, G. (2000). Task-switching, stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35-70). Cambridge, MA: MIT Press.

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological Science,* 15 (9), 610-615.

Arrington, C. M., Logan, G. D., & Schneider, D. W. (2007). Separating cue encoding from target processing in the explicit task-cuing procedure: Are there “true” task switch effects? *Journal of Experimental Psychology: Learning, Memory and Cognition,* 33 (3), 484-502.

Ashby, F. G., Ell, S. W., & Waldron, E. M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition,* 31, 1114-1125.

Bissett, P. G., & Logan, G. D. (2011). Post-stop-signal slowing: Strategies dominate reflexes and implicit learning. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 746–757.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.

Chein, J. M., & Schneider, W. (2012). The brain’s learning and control architecture. *Current Directions in Psychological Science*, 21 (2), 78-84.

Cohen-Kdoshay, O., & Meiran, N. (2007). The representation of instructions in working memory leads to autonomous response activation: Evidence from the first trials in the flanker paradigm. *Quartely Journal of Experimental Psychology,* 60 (8), 1140-1154.

Cohen-Kdoshay, O., & Meiran, N. (2009). The representation of instructions operates like a prepared reflex. *Experimental Psychology,* 56 (2), 128-133.

Cole, M. W., Bagic, A., Kass, R., & Schneider, W. (2010). Prefrontal dynamics in rapid instructed task learning reverse with practice. *The Journal of Neuroscience*, 30 (42), 14245-14254.

Collins, A. G. E., & Frank, M. J. (2013). Cognitive control over learning: Creating, clustering and generalizing task-set structure. *Psychological Review*, 120 (1), 190-229.

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, 47 (3), 153-176.

Denkinger, B., & Koutstaal, W. (2009). Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35 (3), 742-756.

Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science,* 21 (4), 227-231.

Duthoo, W., Abrahamse, E. L., Braem, S., Boehler, C. N., & Notebaert, W. (2014). The heterogeneous world of congruency sequence effects: An update. *Frontiers in Psychology*, 5, 1-9.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences*, 12 (10), 374-380.

Frings, C., Moeller, B., & Rothermund, K. (2013). Retrieval of event files can be conceptually mediated. *Attention, Perception and Psychophysiology*, 75, 700-709.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, 121 (4), 480-506.

Hall, G. (2002). Associative structures in Pavlovian and instrumental conditioning. In C. R. Gallistel (Ed.) *Stevens Handbook of Experimental Psychology* (pp. 1-45). New York: John Wiley & Sons.

Hazeltine, E., & Schumacher, E. H. (2016). Understanding central processes: The case against simple stimulus-response associations and for complex task representation. In B. H. Ross (Ed.) *Psychology of Learning and Motivation* (pp. 195-245). New York: Academic Press.

Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus-response bindings in priming. *Trends in Cognitive Sciences*, 18 (7), 376-384.

Homa, D., Cross, J., Cornell, D., Goldman, D., & Shwartz, S. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *Journal of Experimental Psychology*, 101 (1), 116-122.

Horner, A. J., & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal of experimental Psychology: Learning, Memory, and Cognition*, 35 (3), 757-779.

Horner, A. J., & Henson, R. N. (2011). Stimulus-response bindings code both abstract and specific representations of stimuli: evidence from a classification priming design that reverses multiple levels of response representation. *Memory and Cognition*, 39, 1457-1471.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Phillip, A., & Koch, I. (2010). Control and interference in task switching - A review. *Psychological Bulletin,* 136, 849-874.

Kramer, A. F., Strayer, D. L., & Buckley, J. (1990). Development and transfer of automatic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 16 (3), 505-522.

Kramer, A. F., Strayer, D. L., & Buckley, J. (1991). Task versus component consistency in the development of automatic processing: A psychophysiological assessment. *Psychophysiology*, 28 (4), 425-437.

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science*, 8 (2), 225-247.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, 4: 863, 1-12.

Liefooghe, B., De Houwer, J., & Wenke, D. (2013). Instruction-based response activation depends on task preparation. *Psychonomic Bulletin and Review,* 20, 481-487.

Liefooghe, B., Wenke, D., & De Houwer, J. (2012). Instruction-based task-rule congruency effects. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 38 (5), 1325-1335.

Logan, G. D. (1988). Toward and instance theory of automatization. *Psychological Review*, 95 (4), 492-527.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, 22, 1-35.

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance,* 29, 575-599.

Logan, G. D., & Schneider, D. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (3), 347-363.

Maddox, T. W., Glass, B. D., O'Brien, J. B., Filoteo, J. V., & Ashby F. G. (2010). Category label and response location shifts in category learning. *Psychological Research*, 74, 219-236.

Mayr, U., & Bryck, R. L. (2005). Sticky rules: Integration between abstract rules and specific actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31 (2), 337-350.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85 (3), 207-238.

Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015a). The power of instructions: Proactive configuration of stimulus-response translation. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 41 (3), 768-786.

Meiran, N., Pereg, M., Kessler, Y., Cole, M. W., & Braver, T. S. (2015b). Reflexive activation of newly instructed stimulus-response rules: Evidence from lateralized readiness potentials in no-go trials. *Cognitive, Affective and Behavioral Neuroscience, I15 (2), 365-373.*

Monsell, S. (2003). Task switching. *Trends in Cognitive Science*, 7 (3), 134-140.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). *BayesFactor: Computation of Bayes factors for common designs (Version 0.9.11-1).*

Moutsopoulou, K., & Waszak, F. (2012). Across-task priming revisited: Response and task conflicts disentangled using ex-Gaussian distribution analysis. *Journal of Experimental Psychology: Human Perception & Performance,* 38 (2), 367-374.

Moutsopoulou, K., & Waszak, F. (2013). Durability of classification and action learning: Differences revealed using ex-Gaussian distribution analysis. *Experimental Brain Research,* 226 (3), 373-382.

Moutsopoulou, K., Yang, Q., Desantis, A., & Waszak, F. (2015). Stimulus-classification and stimulus-action associations: Effects of repetition learning and durability. *Quarterly Journal of Experimental Psychology,* 68 (9), 1744-1757.

Murphy, G. L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Neisser, U., & Beller, H. K. (1965). Searching through word lists. *British Journal of Psychology*, 56 (4), 349-358.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, 115 (1), 39-57.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104 (2), 266-300.

Nosofsky, R. M., Stanton, R. D., & Zaki, S. R. (2005). Procedural interference in perceptual classification: Implicit learning or cognitive complexity? *Memory & Cognition*, 33, 1256-1271.

Pereg, M., & Meiran, N. (2017). Evidence for instructions-based updating of task-set representations: the informed fadeout effect. *Psychological Research.* Advance online publication. doi: 10.1007/s00426-017-0842-1

Pashler, H., & Baylis, G. (1991). Procedural learning: 1. Locus of practice effects in speeded choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17 (1), 20-32.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of experimental psychology*, 77 (3), 353-363.

Pothos, E.M. and Wills, A.J. (Eds.) (2011). *Formal approaches in categorization.* Cambridge, England: Cambridge University Press.

R Development Core Team, (2015). *R: A language and environment for statistical computing.* Vienna, Austria.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382-407.

Ruge, H., & Wolfensteller, U. (2010). Rapid formation of pragmatic rule representations in the human brain during instruction-based learning. *Cerebral Cortex*, 20, 1654-1667.

Schmidt, J. R., & Liefooghe, B. (2016). Feature integration and task switching: Diminished switch costs after controlling for stimulus, response, and cue repetitions. *PLoS ONE*, 11 (3).

Schneider, W., & Fisk, A. D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10 (1), 1-15.

Scheider, W., & Shiffrin, R. F. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review,* 84 (1), 1-66.

Schnyer, D. M., Dobbins, I. G., Nichols, L., Davis, S., Verfaellie, M., & Schacter, D. L. (2007). Item to decision mapping in rapid response learning. *Memory and Cognition*, 35 (6), 1472-1482.

Smith, J. D., & Minda, J. P. (1998).Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (6), 1411-1436.

Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28 (4), 800-811.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, 136, 601-626.

Verbruggen, F., Best, M., Bowditch, W. A., Stevens, T., & McLaren, I. P. L. (2014). The inhibitory control reflex. *Neuropsychologia*, 65, 263-278.

Verbruggen, F., & Logan, G. D. (2008). Automatic and controlled response inhibition: associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, 137 (4), 649-672.

Verbruggen, F., Logan, G. D., Liefooghe, B., & Vandierendonck, A. (2008). Short-term aftereffects of response inhibition: repetition priming or between-trial control adjustments? *Journal of Experimental Psychology: Human Perception and Performance*, 34 (2), 413-426.

Waszak, F., Hommel, B., & Allport, A. (2004). Semantic generalization of stimulus-task bindings. *Psychonomic Bulletin and Review*, 11 (6), 1027-1033.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 *t* tests. *Perspectives in Psychological Science*, 6 (3), 291-298.

Wills, A.J. (2013). Models of categorization. In D. Reisberg (Ed.) *Oxford Handbook of Cognitive Psychology* (pp. 346-357). Oxford: Oxford University Press.

Wills, A. J. & McLaren, I. P. L. (1998). Perceptual learning and free classification. *Quarterly Journal of Experimental Psychology, 51B,* 33-58.

Wills, A. J., Noury, M., Moberly, N. J., & Newport, M. (2006). Formation of category representations. *Memory and Cognition*, 34 (1), 17-27.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs.” *Psychological Research/Psychologische Forschung,* 63, 212-233.

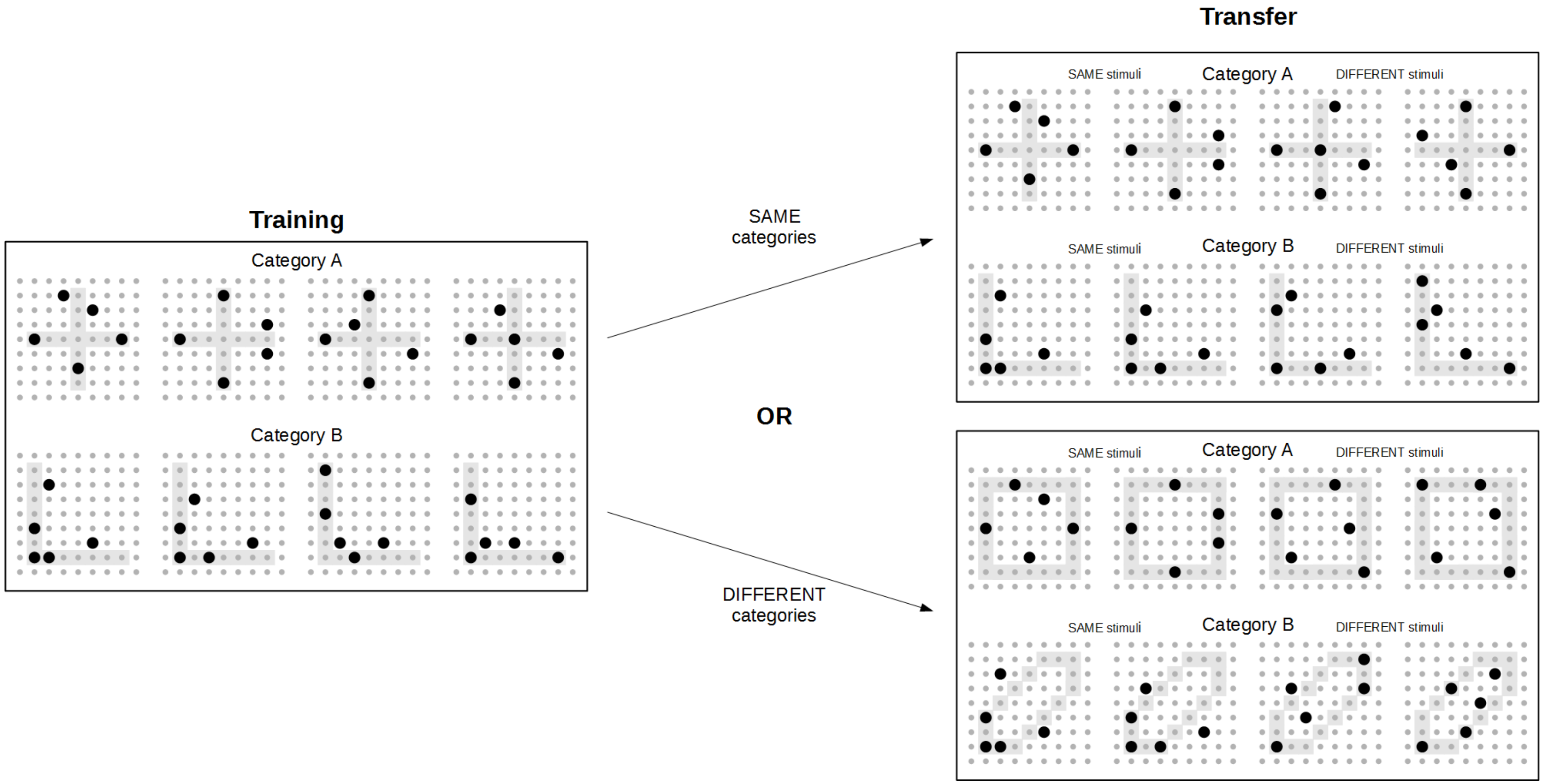

Figure 1
Overview of the experimental design and some example stimuli from each condition. Novel stimuli and categories were introduced for each training block. At transfer the categories could either be novel or repeated from training. Whether the transfer categories were novel or repeated, half of the transfer stimuli were novel and half were repeated from the preceding training block.

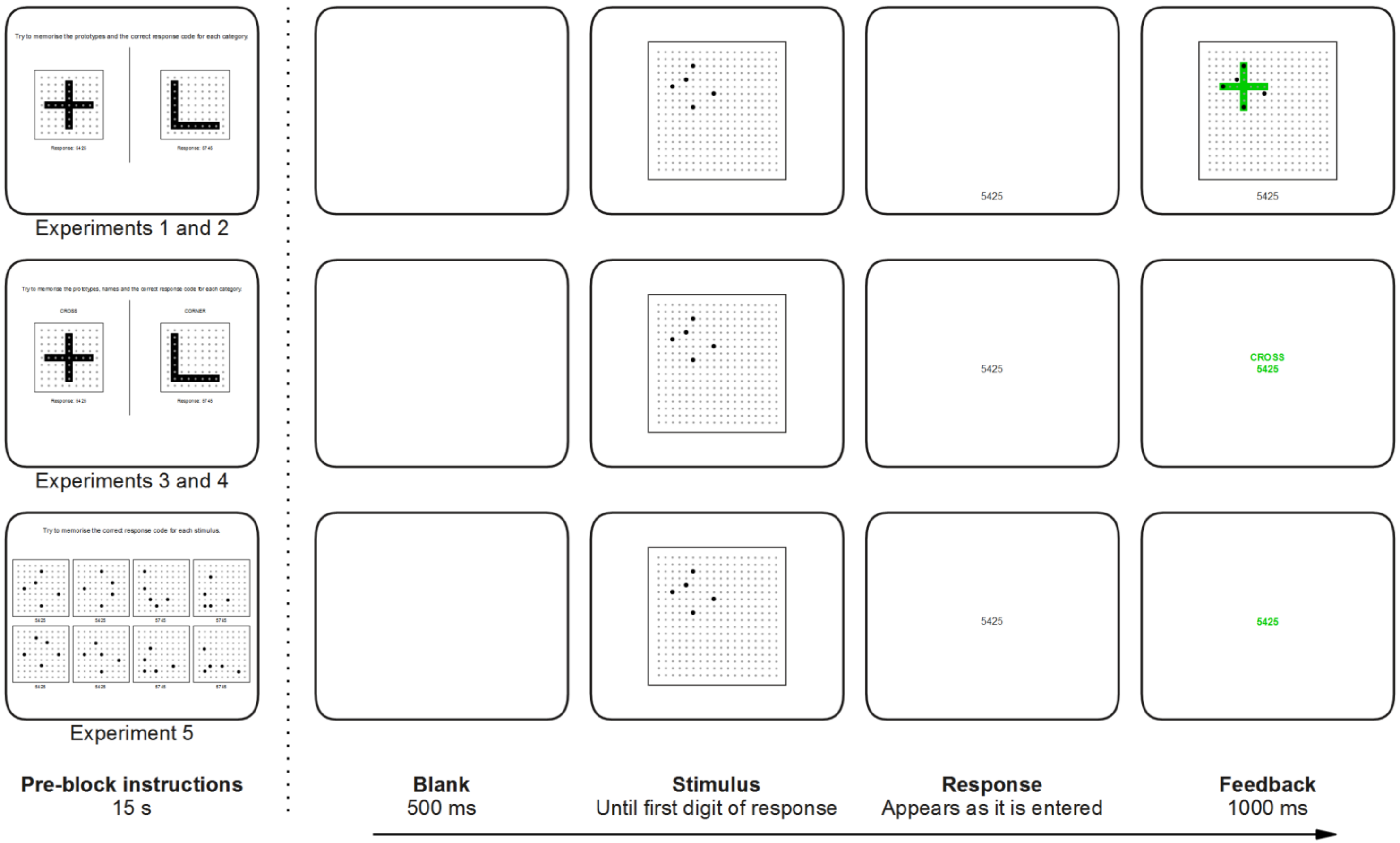

Figure 2
Pre-block instructions screen and the timeline of a single trial for Experiments 1 and 2 (top), Experiments 3 and 4 (middle) and Experiment 5 (bottom). In all experiments the immediate feedback was presented in either green (correct) or red (error).

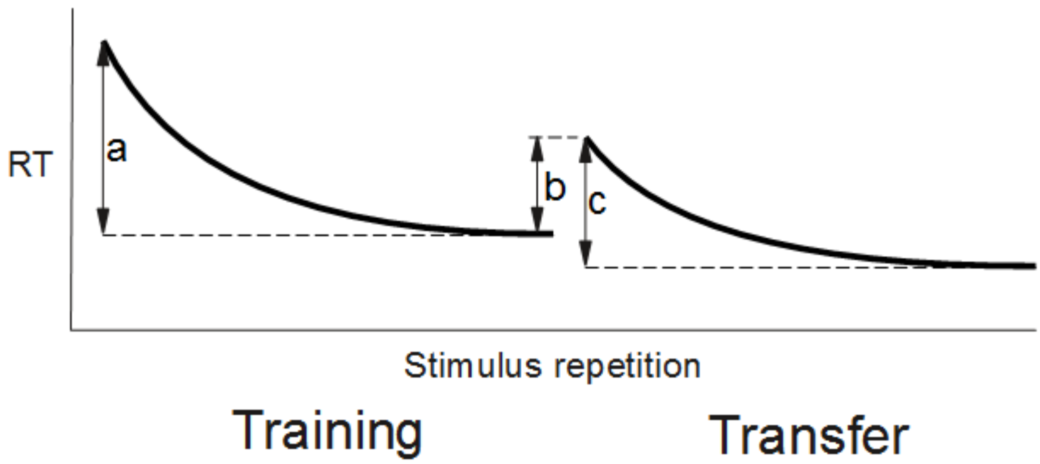

Figure 3
Graphical representation of the three measures used for statistical analysis: a = ‘training learning’; b = ‘transfer effect’; c = ‘transfer learning’ (see text for full description).

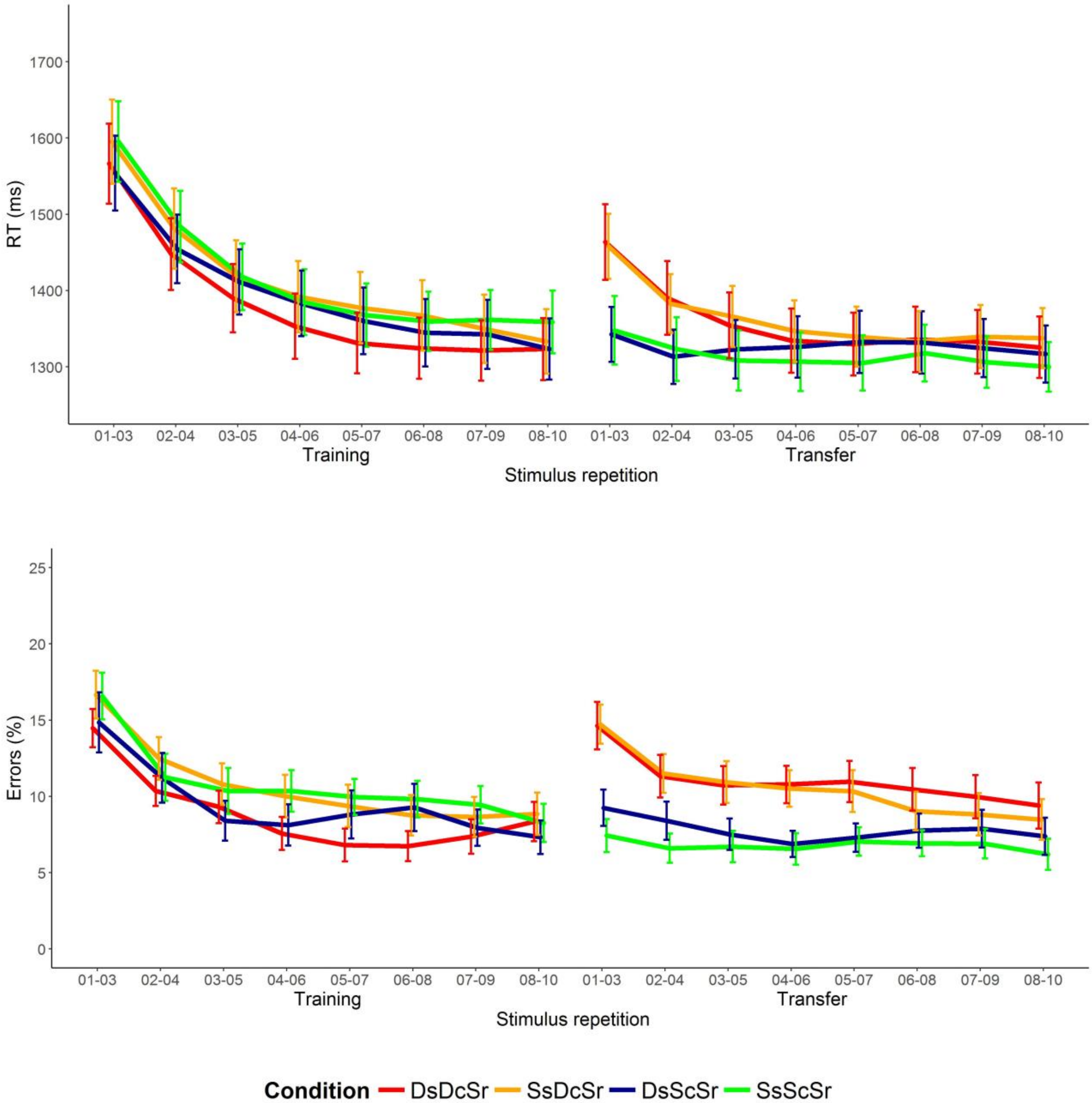

Figure 4
Mean RTs (top) and errors (bottom) from the training and transfer blocks in Experiment 1 plotted as a function of condition and stimulus repetition. Error bars show the standard error of the mean. (See text for condition coding).

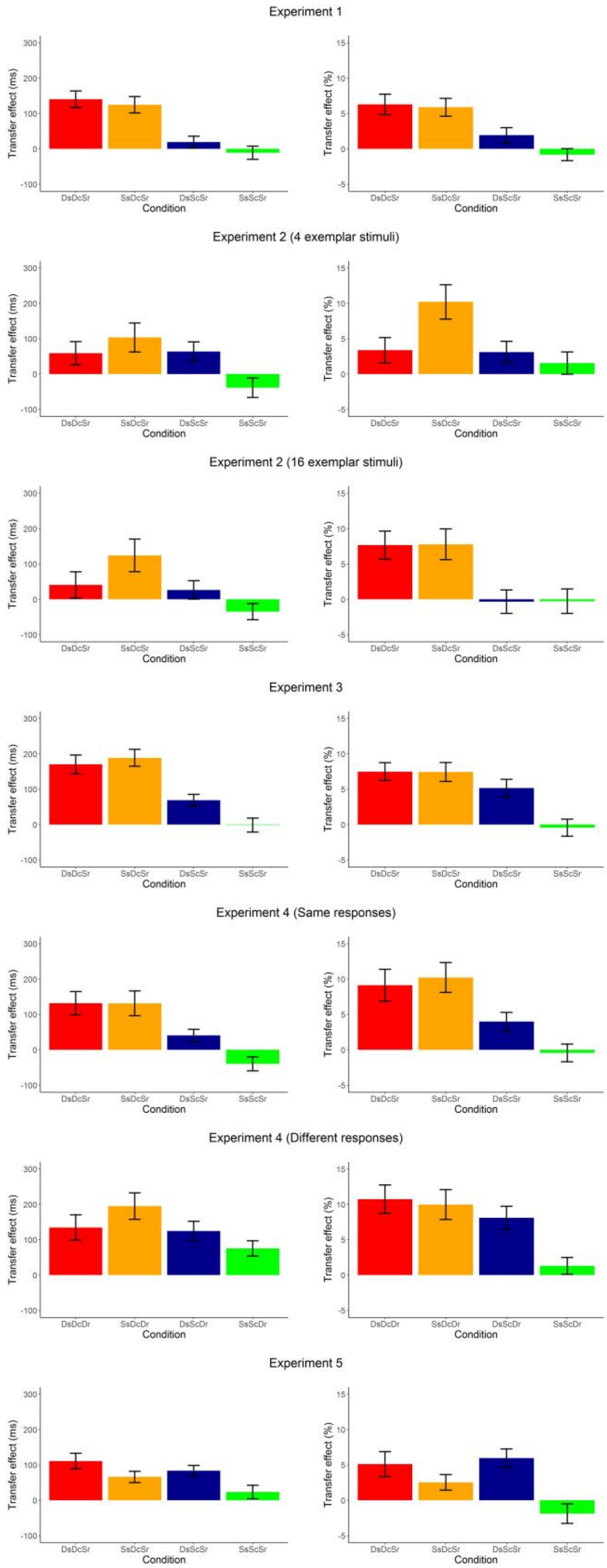

Figure 5
Mean 'transfer effect' for RTs (left) and proportion of errors (right) from each experiment as a function of condition. Error bars show the standard error of the mean. (See text for condition coding).

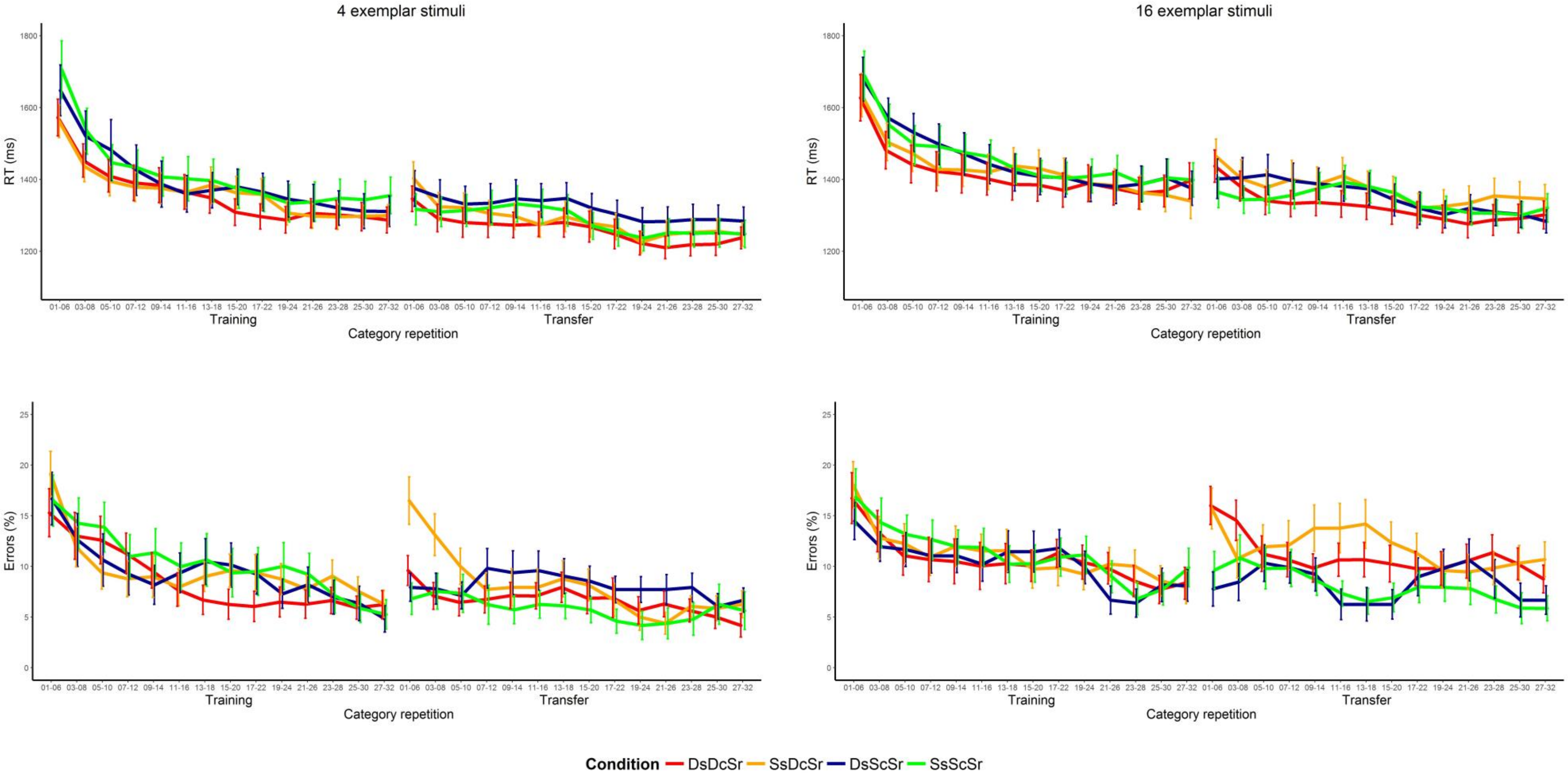

Figure 6
Mean RTs (top) and errors (bottom) from the training and transfer blocks in Experiment 2 where the categories included 4 (left) or 16 (right) exemplar stimuli plotted as a function of condition and category repetition. Error bars show the standard error of the mean. (See text for condition coding).

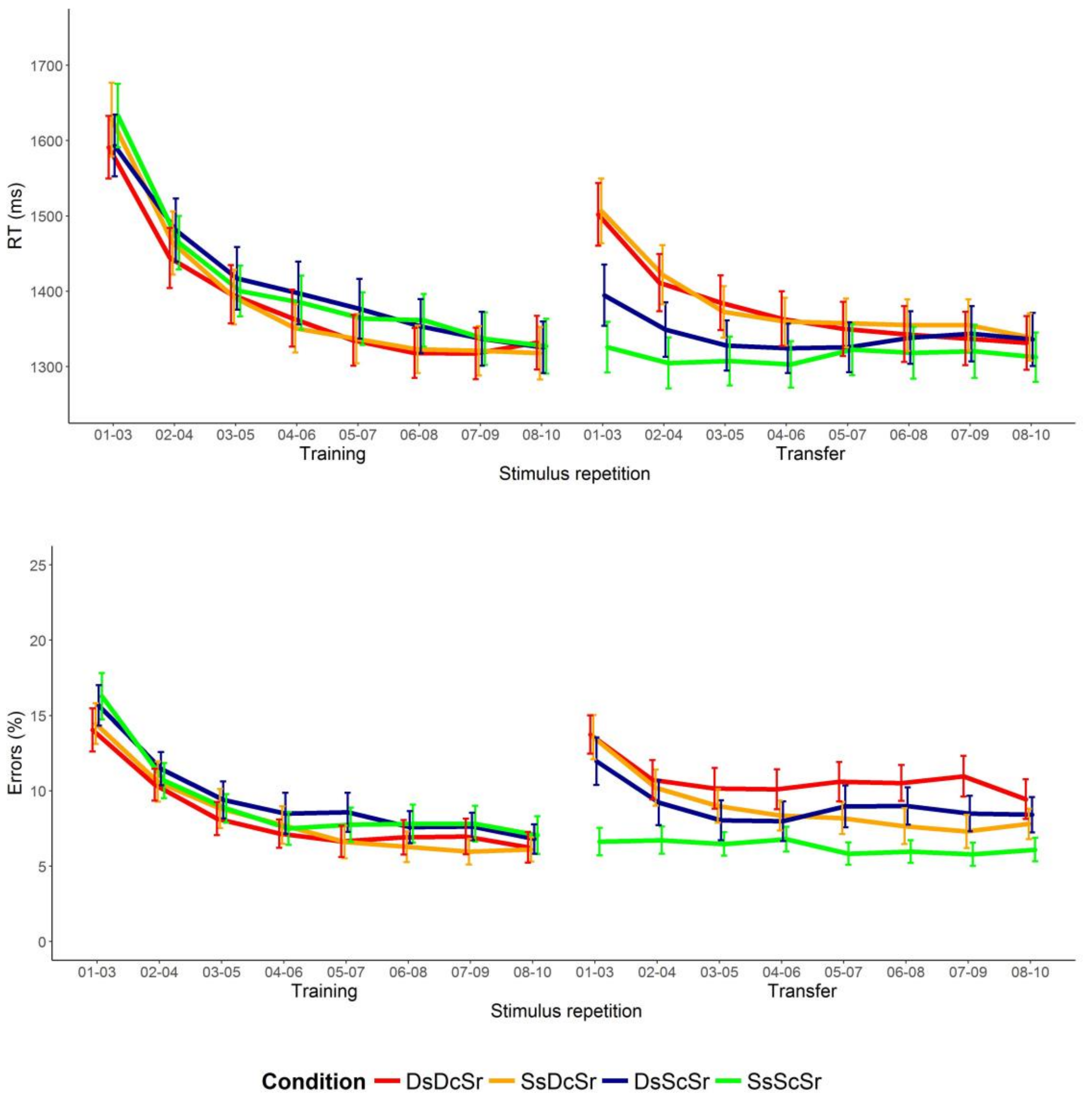

Figure 7

Mean RTs (top) and errors (bottom) from the training and transfer blocks in Experiment 3 plotted as a function of condition and stimulus repetition. Error bars show the standard error of the mean. (See text for condition coding).

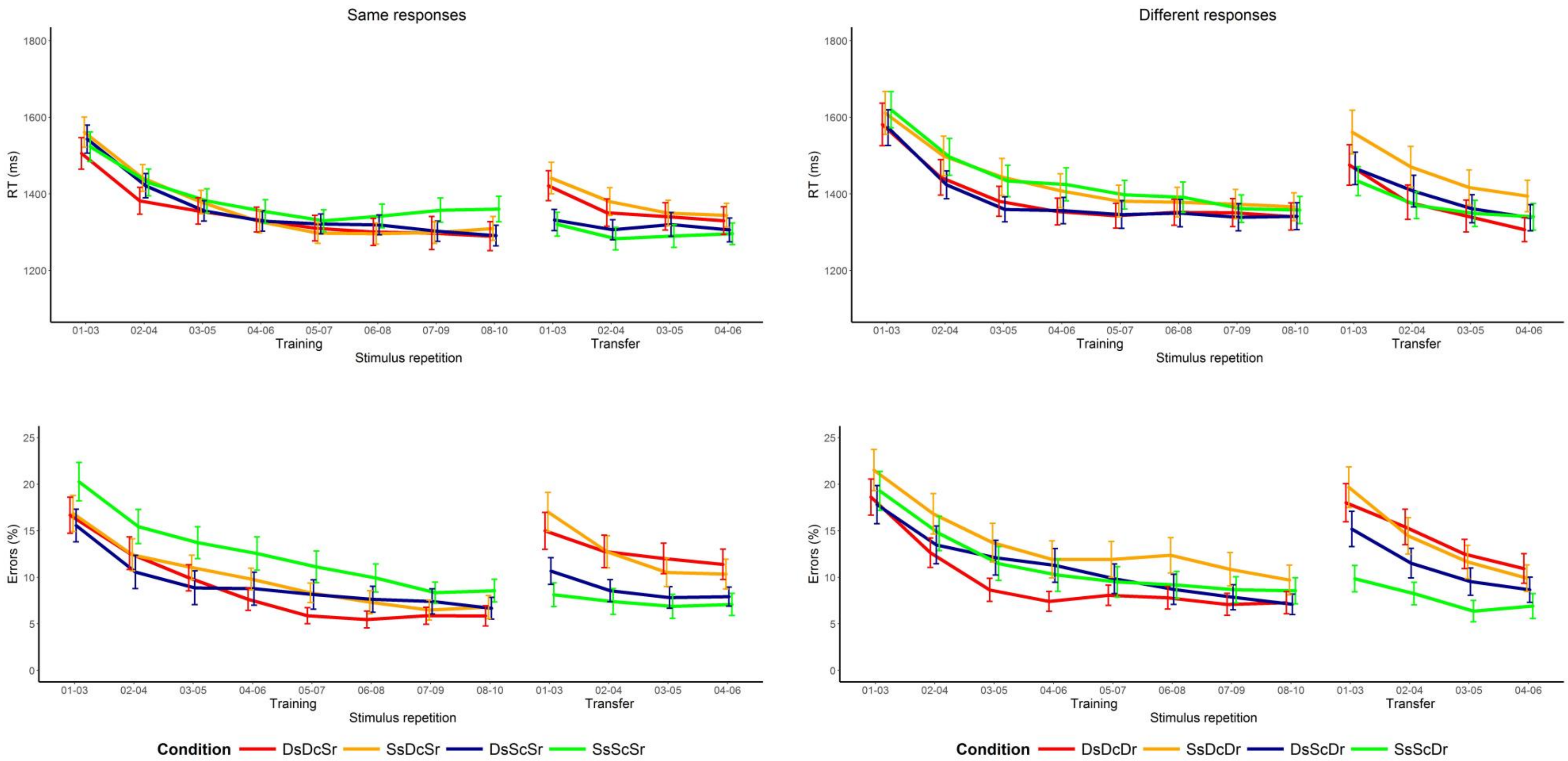

Figure 8

Mean RTs (top) and errors (bottom) from the training and transfer blocks in Experiment 4 where the responses used at transfer were the same as those used during training (left) or were novel (right) plotted as a function of condition and stimulus repetition. Error bars show the standard error of the mean. (See text for condition coding).

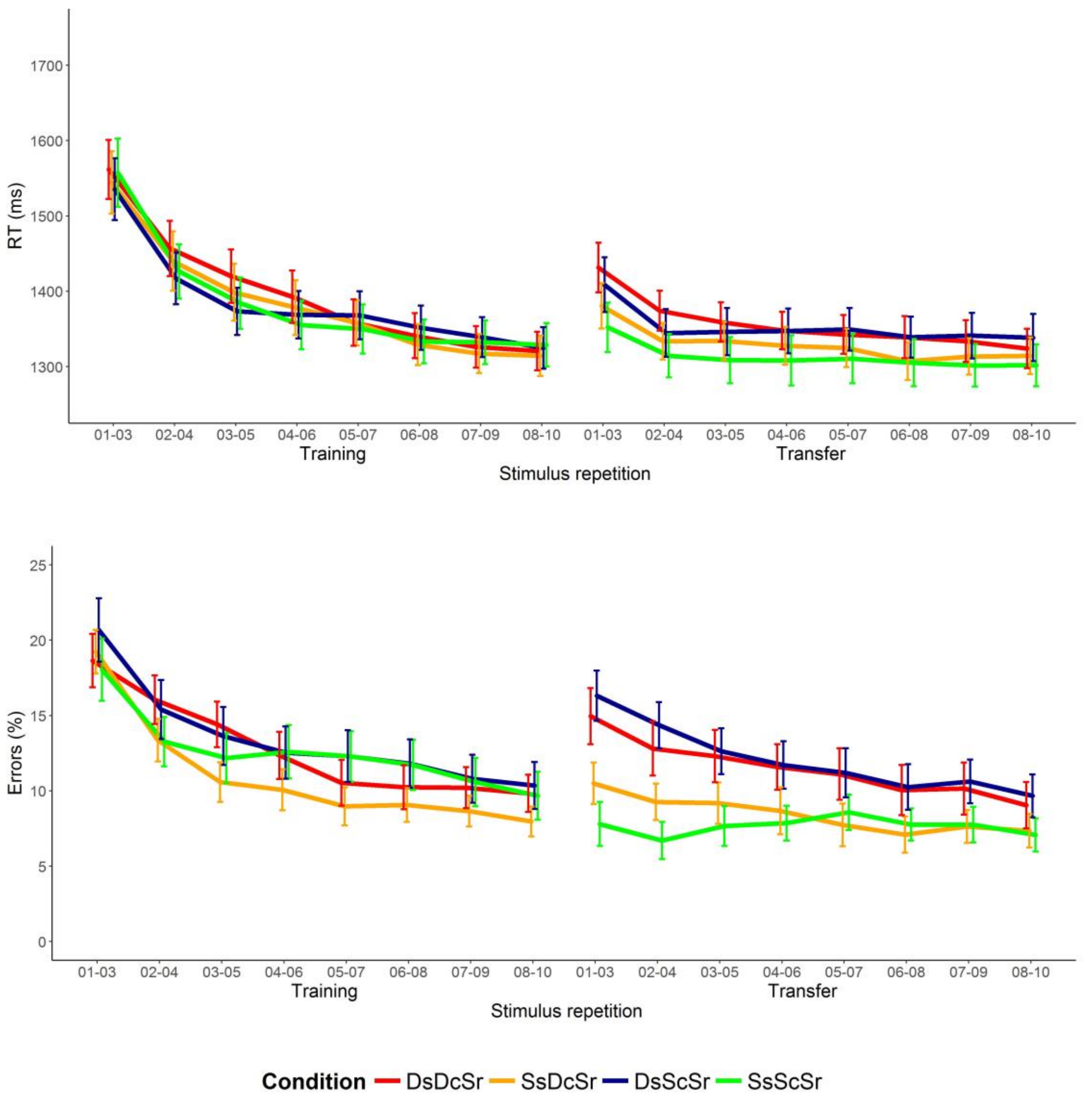

Figure 9
Mean RTs (top) and errors (bottom) from the training and transfer blocks in Experiment 5 plotted as a function of condition and stimulus repetition. Error bars show the standard error of the mean. (See text for condition coding).

Table 1: ANOVA Results from Experiment 1. Equivalent Bayes Factors are also Reported.

Training Learning

	RT						**Errors**					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	13857.39	0.45	0.506	0.004	4.556±5.2%	(1, 39)	62.02	0.98	0.327	0.008	3.543±4.7%
Category	(1, 39)	16176.88	0.87	0.356	0.009	3.480±5.2%	(1, 39)	59.75	0.61	0.439	0.005	4.279±4.6%
Stimulus*Category	(1, 39)	12188.24	0.13	0.720	0.001	3.896±5.3%	(1, 39)	66.91	0.12	0.734	0.001	3.920±4.2%

Transfer Effect

	RT						**Errors**					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	6420.92	3.26	0.079	0.015	2.526±4.6%	(1, 39)	51.20	1.96	0.169	0.016	1.733±19.8%
Category	(1, 39)	14264.07	46.18	<0.001	0.325	<0.001±5.7%	(1, 39)	67.38	18.17	<0.001	0.168	<0.001±20.0%
Stimulus*Category	(1, 39)	14424.68	0.14	0.709	0.001	5.142±16.7%	(1, 39)	36.90	1.49	0.230	0.009	2.026±19.8%

Transfer Learning

	RT						**Errors**					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	5323.67	0.03	0.854	<0.001	5.882±2.6%	(1, 39)	31.77	0.04	0.835	<0.001	8.434±28.0%
Category	(1, 39)	23899.84	14.30	0.001	0.175	<0.001±14.0%	(1, 39)	70.98	9.96	0.003	0.116	0.003±4.9%
Stimulus*Category	(1, 39)	12104.04	1.32	0.258	0.010	2.679±3.7%	(1, 39)	34.88	0.79	0.381	0.005	3.188±2.9%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Table 2: ANOVA results from Experiment 2. Equivalent Bayes Factors are also Reported.

Training Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	61586.84	0.14	0.711	0.001	4.925±40.6%	(1, 39)	135.39	1.09	0.303	0.003	5.721±10.5%
Category	(1, 39)	80079.47	2.71	0.108	0.013	0.967±40.8%	(1, 39)	338.71	0.24	0.628	0.001	6.142±8.5%
Exemplar	(1, 39)	100071.08	0.65	0.425	0.004	3.693±41.5%	(1, 39)	189.28	4.61	0.038	0.015	0.954±8.7%
Stimulus*Category	(1, 39)	37537.13	0.21	0.647	<0.001	3.423±40.7%	(1, 39)	219.16	0.70	0.408	0.003	4.055±9.2%
Stimulus*Exemplar	(1, 39)	31431.00	0.59	0.449	0.001	3.204±40.8%	(1, 39)	167.97	0.02	0.886	<0.001	5.619±8.5%
Category*Exemplar	(1, 39)	87094.37	0.29	0.594	0.001	3.247±41.1%	(1, 39)	213.99	0.95	0.335	0.004	3.365±8.5%
Stimulus*Category*Exemplar	(1, 39)	38804.74	1.64	0.208	0.004	1.694±40.8%	(1, 39)	187.12	0.29	0.597	0.001	3.648±9.1%

Transfer Effect

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	27333.30	0.23	0.635	0.001	6.299±15.0%	(1, 39)	88.10	1.70	0.200	0.004	5.416±18.3%
Category	(1, 39)	48991.39	9.87	0.003	0.044	0.016±17.5%	(1, 39)	168.92	18.44	<0.001	0.076	<0.001±7.0%
Exemplar	(1, 39)	71234.42	0.07	0.799	<0.001	6.706±15.0%	(1, 39)	204.10	0.28	0.599	0.002	6.590±8.4%
Stimulus*Category	(1, 39)	31474.12	13.53	0.001	0.039	0.028±15.5%	(1, 39)	103.69	3.42	0.072	0.009	1.914±12.3%
Stimulus*Exemplar	(1, 39)	24461.58	1.30	0.261	0.003	3.226±14.8%	(1, 39)	121.17	1.08	0.306	0.003	3.662±7.6%
Category*Exemplar	(1, 39)	51349.84	0.12	0.729	0.001	4.830±15.1%	(1, 39)	162.00	1.57	0.218	0.007	2.254±9.8%
Stimulus*Category*Exemplar	(1, 39)	15530.06	<0.01	0.970	<0.001	4.885±17.1%	(1, 39)	119.57	2.91	0.096	0.009	1.575±15.6%

Transfer Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	27335.70	0.84	0.365	0.002	6.936±6.8%	(1, 39)	96.20	1.41	0.242	0.004	6.732±28.1%
Category	(1, 39)	84463.84	2.32	0.135	0.019	0.624±9.1%	(1, 39)	174.25	12.24	0.001	0.055	0.003±14.4%
Exemplar	(1, 39)	48962.17	0.01	0.929	<0.001	9.084±7.7%	(1, 39)	98.72	0.04	0.835	<0.001	7.712±6.0%
Stimulus*Category	(1, 39)	25617.60	3.18	0.082	0.008	2.460±9.4%	(1, 39)	144.33	<0.01	0.999	<0.001	6.225±8.3%
Stimulus*Exemplar	(1, 39)	23363.60	2.98	0.092	0.007	2.475±6.0%	(1, 39)	91.97	1.01	0.322	0.003	5.075±15.6%
Category*Exemplar	(1, 39)	40291.63	0.02	0.879	<0.001	6.367±6.1%	(1, 39)	178.24	0.95	0.336	0.005	3.345±12.1%
Stimulus*Category*Exemplar	(1, 39)	13722.35	0.05	0.829	<0.001	5.281±9.1%	(1, 39)	152.45	3.15	0.084	0.013	0.848±9.2%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Table 3: ANOVA Results from Experiment 3. Equivalent Bayes Factors are also Reported.

Training Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	13444.58	5.90	0.020	0.035	0.833±10.4%	(1, 39)	34.55	0.24	0.626	0.001	5.402±9.2%
Category	(1, 39)	31951.33	0.01	0.937	<0.001	5.280±9.2%	(1, 39)	69.99	0.55	0.462	0.007	4.843±20.6%
Stimulus*Category	(1, 39)	11114.51	0.14	0.710	0.001	4.051±9.8%	(1, 39)	43.21	0.01	0.923	<0.001	4.008±8.1%

Transfer Effect

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	11455.78	2.33	0.135	0.013	2.976±2.7%	(1, 39)	60.36	5.27	0.027	0.042	0.427±4.3%
Category	(1, 39)	29659.01	28.61	<0.001	0.291	<0.001±2.7%	(1, 39)	69.30	14.98	<0.001	0.126	0.002±4.5%
Stimulus*Category	(1, 39)	11777.93	6.73	0.013	0.037	0.550±2.4%	(1, 39)	55.05	5.62	0.023	0.041	0.384±4.4%

Transfer Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	13257.90	1.76	0.193	0.013	2.750±4.3%	(1, 39)	47.32	0.52	0.477	0.004	4.569±4.7%
Category	(1, 39)	22514.05	31.46	<0.001	0.291	<0.001±7.1%	(1, 39)	67.73	5.23	0.028	0.052	0.178±4.2%
Stimulus*Category	(1, 39)	8449.87	2.13	0.153	0.010	2.497±4.5%	(1, 39)	52.05	3.89	0.056	0.030	0.615±4.0%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Table 4: ANOVA Results for the Merged Data from Experiments 1-3. Equivalent Bayes Factors are also Reported.

	Transfer Effect											
	RT						**Errors**					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 119)	11954.14	6.70	0.011	0.011	1.622±3.1%	(1, 119)	76.62	0.54	0.464	0.001	8.327±3.0%
Category	(1, 119)	33342.45	47.01	<0.001	0.176	<0.001±2.7%	(1, 119)	107.41	28.25	<0.001	0.092	<0.001±2.5%
Stimulus*Category	(1, 119)	16427.82	12.71	0.001	0.028	0.063±3.1%	(1, 119)	67.45	13.11	<0.001	0.029	0.040±2.6%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Table 5: ANOVA Results from Experiment 4. Equivalent Bayes Factors are also Reported.

Training Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	10082.54	0.18	0.674	<0.001	7.677±12.7%	(1, 39)	95.14	0.34	0.563	0.001	8.099±6.6%
Category	(1, 39)	28939.23	0.35	0.559	0.002	5.861±12.1%	(1, 39)	150.60	0.13	0.717	0.001	8.089±5.8%
Response	(1, 39)	42362.19	1.09	0.304	0.007	2.965±13.0%	(1, 39)	158.06	0.31	0.579	0.002	7.316±7.1%
Stimulus*Category	(1, 39)	18613.78	2.60	0.115	0.008	2.106±12.5%	(1, 39)	94.51	0.52	0.477	0.002	5.303±6.5%
Stimulus*Response	(1, 39)	13960.26	2.82	0.101	0.006	2.486±12.8%	(1, 39)	80.37	0.15	0.698	<0.001	6.717±9.1%
Category*Response	(1, 39)	25762.62	0.72	0.401	0.003	4.028±13.4%	(1, 39)	93.50	0.11	0.742	<0.001	6.035±5.7%
Stimulus*Category*Response	(1, 39)	18769.88	6.05	0.018	0.018	0.464±13.2%	(1, 39)	85.83	0.94	0.337	0.003	3.374±6.1%

Transfer Effect

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	11925.26	2.02	0.163	0.003	7.900±26.0%	(1, 39)	100.92	5.89	0.020	0.017	0.763±5.7%
Category	(1, 39)	74193.68	10.33	0.003	0.088	<0.001±4.7%	(1, 39)	85.58	42.98	<0.001	0.096	<0.001±7.2%
Response	(1, 39)	34404.26	10.23	0.003	0.042	0.025±4.9%	(1, 39)	181.09	1.42	0.241	0.007	3.136±6.7%
Stimulus*Category	(1, 39)	17013.67	10.62	0.002	0.022	0.338±5.6%	(1, 39)	103.72	6.44	0.015	0.019	0.510±12.7%
Stimulus*Response	(1, 39)	9812.24	4.26	0.046	0.005	3.217±6.2%	(1, 39)	168.43	0.52	0.474	0.003	4.630±7.1%
Category*Response	(1, 39)	37178.52	2.36	0.133	0.011	1.604±8.4%	(1, 39)	167.18	0.60	0.443	0.003	4.895±11.0%
Stimulus*Category*Response	(1, 39)	19696.23	0.23	0.637	0.001	4.367±6.1%	(1, 39)	78.05	0.02	0.900	<0.001	4.488±5.1%

Transfer Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	10753.52	0.50	0.483	0.001	10.166±40.1%	(1, 39)	89.62	0.01	0.907	<0.001	8.619±7.7%
Category	(1, 39)	30415.30	10.46	0.002	0.067	0.001±13.3%	(1, 39)	60.12	15.97	<0.001	0.041	0.029±9.5%
Response	(1, 39)	25683.26	19.68	<0.001	0.102	<0.001±10.2%	(1, 39)	90.33	8.28	0.006	0.032	0.098±9.6%
Stimulus*Category	(1, 39)	12375.44	0.67	0.417	0.002	4.235±10.8%	(1, 39)	81.03	7.62	0.009	0.027	0.190±10.4%
Stimulus*Response	(1, 39)	8337.45	1.08	0.305	0.002	4.106±10.3%	(1, 39)	82.26	0.31	0.579	0.001	5.308±6.7%
Category*Response	(1, 39)	17291.04	0.16	0.696	0.001	5.855±18.2%	(1, 39)	81.64	0.05	0.827	<0.001	6.026±6.5%
Stimulus*Category*Response	(1, 39)	9224.45	0.39	0.534	0.001	4.085±13.4%	(1, 39)	94.50	0.12	0.731	0.001	4.180±6.1%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Table 6: ANOVA Results from Experiment 5. Equivalent Bayes Factors are also Reported.

Training Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	12569.77	0.04	0.844	<0.001	6.005±4.8%	(1, 39)	58.68	0.04	0.839	<0.001	5.805±4.4%
Category	(1, 39)	17585.04	0.59	0.448	0.007	3.996±4.6%	(1, 39)	88.44	0.22	0.645	0.002	5.438±6.7%
Stimulus*Category	(1, 39)	9613.31	0.84	0.365	0.005	3.644±8.9%	(1, 39)	55.40	3.54	0.067	0.024	1.087±4.4%

Transfer Effect

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	7631.24	14.44	<0.001	0.076	0.060±2.7%	(1, 39)	54.25	20.10	<0.001	0.105	0.006±3.2%
Category	(1, 39)	17865.37	2.74	0.106	0.035	0.798±5.8%	(1, 39)	104.83	1.22	0.277	0.014	2.537±4.3%
Stimulus*Category	(1, 39)	8756.84	0.29	0.595	0.002	3.897±2.6%	(1, 39)	78.84	3.50	0.069	0.029	0.907±12.1%

Transfer Learning

	RT						Errors					
Effect	***DF***	***MSE***	***F***	***p***	***η2***	***BF***	***DF***	***MSE***	***F***	***p***	***η2***	***BF***
Stimulus	(1, 39)	9863.75	3.76	0.060	0.025	1.163±16.9%	(1, 39)	38.46	19.65	<0.001	0.128	0.003±3.1%
Category	(1, 39)	18593.30	1.52	0.225	0.019	1.640±16.9%	(1, 39)	53.80	0.53	0.472	0.005	4.773±7.1%
Stimulus*Category	(1, 39)	8449.07	0.56	0.459	0.003	2.996±16.9%	(1, 39)	40.15	2.49	0.123	0.019	1.694±4.6%

Note: Bayes factors indicate whether removal of the effect/interaction from the model would materially impair its fit. Thus Bayes factors<1 indicate that the effect/interaction is an important contributor to the model.

Appendix

Templates and names for all categories used in the experiments reported in the manuscript. All categories were used in Experiment 4. Experiments 1, 3, and 5 used the first seven rows. Experiment 2 used only the first five rows. The category names were used to provide feedback to participants during Experiments 3 and 4. Each row in the figure shows the templates for the two categories used during a given training-transfer phase. Templates in the columns marked ‘training’ were used during the training phase and also the transfer phase for conditions where the categories were repeated (e.g., DsScSr, SsScSr). Templates in the columns marked ‘transfer’ were used during the transfer phase for conditions where the categories changed (e.g., SsDcSr, DsDcSr).