Psychological Research (2020) 84:1610–1621
https://doi.org/10.1007/s00426-019-01171-9

ORIGINAL ARTICLE

Task structure boundaries affect response preparation

Savannah L. Cookson[1] · **Eliot Hazeltine**[2] · **Eric H. Schumacher**[3]

Received: 26 July 2018 / Accepted: 23 March 2019 / Published online: 1 April 2019

Abstract

Does cognitive control operate globally (across task sets) or locally (within a task set)? Recently, two of the current co-authors (Hazeltine and Schumacher 2016; Schumacher and Hazeltine 2016) proposed that humans represent tasks as *task files*: hierarchically structured, compartmentalized subsets of our current goals and motivations, task instructions, and relevant stimuli and responses that are selected during task performance according to associated contextual rules. Here, we hypothesize that these task representations bound the implementation of cognitive control at distinct levels of this hierarchical structure. To investigate how task structure influences the implementation of control processes, we conducted a pair of experiments that utilized a precuing procedure. To manipulate task structure, we gave participants mappings in which two stimulus sets were either mapped so that each set was separated by response hand or both sets were interleaved across hands. In Experiment 1, participants responded to sets of images distinguished by their semantic category; in Experiment 2, they responded to sets based on different perceptual features (viz., location or color). During each experiment, precues could give information about the stimulus category or response hand for the upcoming target. The results indicate that participants with separated mappings represented the task hierarchically, while those with interleaved mappings did not. This pattern was consistent across experiments, despite the differences in the way that each set of stimuli influenced representation of the low-level task features. These findings suggest that task structure can be represented hierarchically, and that this structure supports distinct cognitive control processes at different hierarchical levels.

Introduction

Humans are able to select responses to stimuli as a function of current task goals. Typically, response selection is conceived as the translation of stimuli into responses according to learned associations between them, or their stimulus–response (S–R) mapping. However, a simple association of stimuli and responses is an insufficient representational structure to support a variety of findings from the response selection literature. For example, Fitts and Seeger (1953) demonstrated set-level compatibility effects, in which performance depended on correspondence between the spatial layouts of all possible stimuli and responses. Given this and other examples (reviewed by Hazeltine and Schumacher 2016), we proposed an alternative conception of task representation called a task file. In this view, stimuli and responses are linked by a stimulus–response (S–R) mapping structure in a "set" (Schumacher and Hazeltine 2016). These sets can then be hierarchically associated with contextual information (e.g., goal, rules) and control may be implemented to select between sets according to current context. This structure, then, distinguishes between control processes that are involved in "set level" selection and those that influence processing of the specific stimulus and response features of the task.

In the present experiments, we investigated how task structure contributes to the implementation of control. In each experiment, we gave participants one of two different S–R mapping structures, one of which could be represented with a hierarchical task structure with two subgroups, and one of which had an ambiguous structure. We then assessed each group's ability to use information provided by a cue to aid their performance as a function of the type of information given by the cue and the structure of the task

✉ Savannah L. Cookson
savannah.cookson@berkeley.edu

[1] University of California, Berkeley, 132 Barker Hall, Berkeley, CA 94720-3190, USA

[2] University of Iowa, 11 Seashore Hall E, Iowa City, IA 52242, USA

[3] Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332, USA

representation, as well as how switching or repeating the stimulus and response sets influenced performance.

The precuing procedure

Early work manipulating task representation investigated the locus of the benefit of prior information on subsequent performance as it related to information processing models. One common procedure for investigating this question is the precuing paradigm. In precuing tasks, participants are given information prior to stimulus presentation that reduces the number of possible responses without revealing which specific response will be indicated by the target. If participants can utilize this cue information to prepare for the task in some way prior to stimulus presentation, then they will be faster and more accurate to respond to the upcoming target than when no information is presented.

The precuing procedure was first used to investigate the structure of stimulus–response (S–R) processing by examining the order (if any) in which the motor parameters of a response must be specified to observe a performance benefit. Rosenbaum (1980; see also Rosenbaum 1983) tested whether the order in which motor parameters were specified affected the response time (RT). Participants executed arm movements defined by three features: hand (left or right), direction (away from or toward the body), and extent (far or near center). At the start of each trial, a cue could indicate none, any, or all of these features for the upcoming trial. The magnitude of the benefit for the cues depended on which parameters were indicated. Furthermore, the cue benefits stacked, such that providing information for multiple features provided greater benefit than cueing either feature individually. Rosenbaum proposed that participants used information to specify independent dimensions of an upcoming movement, and that the cue benefit resulted from participants not needing to specify as many dimensions at the time of response.

These results spurred debate over the mechanism underlying the cue benefit. Goodman and Kelso (1980) responded to Rosenbaum's (1980) findings by showing that the correspondence between stimuli and responses could be manipulated to eliminate the differences in cueing benefit for different parameters. Specifically, they cued responses with a four-by-two grid of light-emitting diodes that directly mapped to the response structure by lighting subsets of four diodes on a given trial to indicate different response dimensions to be specified. In this task, participants showed a cue benefit that did not depend on the specific parameter that was cued; instead, the benefit depended simply on the number of remaining possible alternative responses, following the Hick/Hyman law (HHL; Hick 1952; Hyman 1953). They claimed that these results indicated that the benefit was not localized to motor parameter specification; instead, cue information reduced uncertainty in the number of upcoming alternatives. Goodman and Kelso argued that this indicated that the cue benefit was localized to response selection, or the translation of the target to a response alternative as a function of the learned mapping between stimuli and responses.

However, later data called into question Goodman and Kelso's (1980) simple uncertainty explanation. Miller (1982; see also Miller 1983) showed that cueing different subsets of response effectors provided different magnitudes of benefits using a similar procedure, suggesting that motor programming was, in fact, the locus of precuing benefits. Miller presented participants with four light stimuli arranged in a horizontal line in the center of the display. These stimuli were mapped in a spatially compatible fashion to four finger responses on a single hand or split between two hands. Cues were presented in which either all four or a subset of two lights was illuminated to indicate a possible subset of upcoming targets. Miller showed that participants maximally benefitted from cues when the left two stimuli were mapped to left hand responses and the right two to right hand responses. He concluded that these differences in the effectiveness of various cue configurations reflected the structure in the response effector system; that is, that certain anatomical effectors (i.e., hand side) could be prepared for motor execution a priori, similar to Rosenbaum's (1980) motor parameter specification hypothesis.

Further complicating the story, Reeve and Proctor (1984) argued against this response preparation hypothesis. Using the same design as Miller (1982), the authors manipulated the S–R mappings by overlapping participants' response hands so that the fingers oriented left to right were the left-middle, right-index, left-index, and right middle fingers, respectively. This removed the spatial correspondence between the stimulus and response sets. The results did indeed show that the reoriented mapping eliminated the hand benefit (later replicated for the other two stimulus sets from Miller's experiments; see Proctor and Reeve 1985). In fact, regardless of how the hands were configured, participants still showed the largest benefit when cues indicated that the stimuli would be presented on one side of the screen or the other. Reeve and Proctor (1984) argued that this was further evidence that the locus of the cue benefit was in response selection; because the benefit manifested along visual boundaries regardless of the response structure, there must be some aspects of stimulus dimensions that can be more easily preprocessed into response sets (viz., left versus right visual field).

Adam, Hommel, and Umilta (2003, 2005) took a Gestalt perspective of these phenomena and proposed the Grouping Model to account for the seemingly conflicting cueing results in this literature. This account focused on structure present in the stimulus displays and responses and the correspondence between them. They hypothesized that task boundaries were flexibly established as a function of: (1) salience at both stimulus- and response-level groupings, and (2) correspondence between these levels such that the two groupings were mapped compatibly with one another. A key aspect of this model was that it focused on the overarching conceptual sets that could be distinguished in the stimuli and responses, rather than the individual S–R pairs.

To evaluate this model, Adam et al. (2003) conducted a series of experiments using the response-cueing paradigm originally introduced by Miller (1982). They manipulated the grouping saliency of both the stimulus and response-levels, as well as the compatibility between these groups. The results showed that the cueing benefit depended on the saliency of both the stimulus and response groupings. This account suggests that cue benefits are not static properties of the response system but were flexibly shaped according to the structure of the task. Importantly, the Grouping Model de-emphasized the traditional serial information-processing model that focused on the sequence of processes underlying translation of stimuli into their associated responses followed by the preparation and execution of the relevant motor actions; instead, it emphasized the biasing of attention toward task-relevant S–R sets.

Task files in selection and preparation

The experiments conducted by Adam et al. (2003, 2005), and many of the studies their work was based on (e.g., Miller 1982, 1983; Reeve and Proctor 1984; Proctor and Reeve 1985), relied on exogenous cues that directly corresponded to the locations of their associated response buttons and appeared in the same locations as the stimuli they cued. That is, participants responded directly to the position of the stimulus, and likewise could prepare responses to the positions of the cues. This task structure lends itself to the grouping model explanation, which requires a direct correspondence between the stimulus and response sets to facilitate a cue benefit. However, Rosenbaum's (1980) original experiment demonstrated the benefits of mnemonic cues, which do not intrinsically carry any information about the stimulus, the response, or their mapping structure. Instead, they carry mnemonic information that must be interpreted in the context of the task. The grouping model does not suggest a mechanism for such benefits.

The task file account proposed by Schumacher and Hazeltine (2016) may offer an account for both exogenous and mnemonic cue benefits. The task file uses a hierarchical task representation to associate contextual selection criteria with subsets of S–R mappings. In the case of Miller's (1982) design, for example, the left and right stimulus/response sets can be selected according to the side of the screen on which they are presented. If the cues on a given trial appear in the two left-most positions, then, participants are able to select the "left side" set. This selection process simultaneously activates the potentially relevant stimuli and responses within the associated set in preparation for the upcoming task. Then, when the target is presented, participants select between the two S–R pairs within that set to make their response. Conversely, if the cues do not appear on one side or the other, then the this set selection process cannot occur prior to target presentation. The result is that, when the target is presented, participants must complete two selection processes—one for the subtask corresponding to the side of presentation, and one for the S–R set within that subtask. This difference in processing demands results in the benefit seen for cues presented on the left or right side.

We previously conducted a functional magnetic resonance imaging (fMRI) study examining the neural substrates of response preparation to mnemonic cues and their associated stimuli and responses (Cookson et al. 2016). Participants were given a precuing task with four face and four building images mapped to the left and right hands such that faces were mapped to one hand and buildings to the other. This design allowed participants to represent the task as two salient subgroups. At the start of each trial, participants were shown a letter cue that indicated either the category of the stimulus (face "F" or place "P"), the response hand (left "L" or right "R"), or neither (noninformative cue "O"). This design could be represented as a hierarchical task in which cue information selected between two sets, defined jointly by their stimulus category and their response hand.

Participants showed a behavioral benefit for informative versus uninformative cues, regardless of the information contained in the informative cues (Cookson et al. 2016). Specifically, reaction times (RT) were faster for informative cues than informative, and did not differ between F, P, L, and R cues. This is consistent with the selection of a set at cue presentation, rather than selecting for specific stimulus or response features. Furthermore, the fMRI data showed that when participants were given an informative stimulus cue, activity in the motor regions contralateral to the response hand for the indicated subset was greater than that in the ipsilateral regions. Additionally, for that indicated subset, the relevant motor regions showed increased connectivity with regions associated with processing the relevant stimulus category (i.e., fusiform gyrus for face images and parahippocampal and lingual gyri for place images) compared to their connectivity with the unassociated regions. This suggests that participants were preparing a network specific to the trial-relevant S–R subset that included stimulus- and response-level processing regions when given a cue that indicated one subset over another. However, there was no manipulation of task structure, so it was not possible to determine what role the representation of the tasks played in the pattern of results.

Current experiments

We report two experiments that use a precuing procedure similar to Cookson, Hazeltine, and Schumacher (2016) to investigate how the task structure affects the implementation of preparatory control processes. In each experiment, there were two stimulus types (e.g., faces and places) and two response sets (i.e., left- and right-hand). To manipulate task structure, two groups of participants were assigned different S–R mappings. One group had stimulus types segregated between the two response sets, and the other group had the stimulus types interleaved across both response sets. At the start of a trial, a cue was presented indicating either the stimulus type, response hand, or no information. We hypothesized that in the segregated, but not the interleaved, mapping, participants could represent the task as two salient subgroups defined by stimulus type and response group, which could in turn be governed by a second level rule structure associating cue information with the relevant subgroup.

The structural manipulation in this design allowed us to directly investigate the role of task structure in S–R processing. Goodman & Kelso (1980) noted that cues indicating the same number of alternative responses can produce different reductions in RT, appearing to violate the HHL. Rosenbaum's (1980) results further demonstrated that these effects can occur when cues do not have any direct perceptual correspondence to stimuli and/or responses. We propose that these findings do not reflect the structure of the motor system, as proposed by Miller (1982), or the structure of sensorimotor mappings, as proposed by Adam et al. (2003, 2005), but rather reflect the task structure—that is, the hierarchical structure that organizes the encoded S–R pairs. If the structure of the task representation affects participants' capacity to establish hierarchical selection rules, the two groups in these experiments should show differential ability to utilize the cue information versus uninformative cue trials. This will demonstrate that S–R processing is mediated by hierarchical representations that incorporate relationships among S–R pairings rather than simply operating on perceptuomotor processes.

Experiment 1

Experiment 1 investigates how participants may benefit differently from cue information as a function of structure implicitly built into the task using a variant of the design described by Cookson, Hazeltine, and Schumacher (2016). In this experiment, participants could be assigned to one of two mapping groups. In one group (the separate mapping, SM), face and place stimuli were mapped separately to left and right hands, as in the previous report; in the other group (interleaved mapping, IM), the face and place mappings were alternated by finger on both hands so that both stimulus types were represented on each hand. In other words, the SM group learned a mapping with a strong hierarchical subgroup structure, while the IM group learned a mapping with an ambiguous structure. We hypothesized that this difference in structure would allow for rule-based selection between two smaller subgroups in the SM group when information was provided a priori but would require the IM group to operate on one flat task.

Methods

Participants

Experiment 1 included 34 participants recruited from the student population at the Georgia Institute of Technology. All participants were at least 18 years old, right-handed, and had normal or corrected-to-normal vision. Participants gave their informed consent before the experiment and were compensated with course credit. All procedures performed in studies involving human participants were in accordance with the ethical standards of the Georgia Institute of Technology Institutional Review Board and with the 1964 Helsinki declaration and its later amendments.

Stimuli and apparatus

Stimuli were presented on a desktop personal computer using E*Prime 2.0 software. Responses were collected using a standard keyboard. Left-hand responses were mapped to the 'a', 's', 'd', and 'f' keys (little to pointer finger) and right-hand responses to the 'h', 'j', 'k', and 'l' keys (pointer to little finger). Text stimuli, including experiment instructions, feedback, fixation markers, and cue stimuli, were presented in white sans serif font. Cue stimuli were the letters 'F', 'P', 'R', 'L', and 'O'; fixation markers were indicated with a '+'. Stimulus images were presented at 100 × 100 pixels. Face images included four images, two male and two female, from the AR Face Database (Martinez and Benavente 1998) presented in color and showing the hair and shoulders; all image subjects are wearing standard t-shirts. Place images included four color images of still scenes. All stimuli were presented centrally on a black background.

Procedure

Participants came to the lab on two separate days spaced no less than one and no more than 7 days apart. Prior to their arrival at the laboratory, each participant was randomly assigned to one of eight mapping alternatives, categorized into one of two groups. In the IM group (17 participants), face and place stimuli were mapped to the responses in an alternating fashion; the category order (i.e., between hands) and stimulus order (i.e., within hand) within category were counterbalanced to yield four different possible mappings. In the SM group (17 participants), face and place images were segregated by hand; images were again counterbalanced to create four possible mappings.

Following the consent process on the first day, and then on each experiment day, the experimenter verbally walked the participant through the task instructions as they were presented on screen. Each time the participant saw a picture, they were to press its corresponding letter key. After being given these instructions, participants were shown their assigned mapping structure and instructed to use this mapping to respond to the stimuli shown on screen. At no point were participants instructed explicitly about the underlying structure of the mapping.

Following this, they completed a sixteen-trial practice block to familiarize themselves with the trial procedure. They then completed two blocks of 48 trials each in which all trials presented only neutral cues to learn their assigned mapping structure. Next, they completed sixteen experimental blocks that alternated between presenting stimulus-type and response-type cues, described below, using a repeating pattern (stimulus–response–response–stimulus). At the beginning of each block, participants were told that a cue presented before the stimulus may give them information about the upcoming target. The instructions then introduced the cue types that would be relevant for that block and told participants what information each cue afforded. On a given experimental block, neutral cues were shown on 33% of trials, and 50% of the remaining trials showed each of the two possible informative cues for that block (i.e., 33% of the total trials per cue).

On each trial, participants were first shown a cue for 1500 ms. Cues consisted of: neutral (‘O’) cues, which gave no information about the upcoming trial; stimulus cues, which indicated whether the upcoming stimulus image would be a face (‘F’) or place (‘P’) image; and response cues, which indicated whether the upcoming response would be located on the left (‘L’) or right (‘R’) hand. Cues were 100% valid. Following the cue, participants saw a fixation cross (“cue-stimulus interval”, CSI) for 500 ms, then were presented the target stimulus for 1500 ms. Participants were instructed to make the response corresponding to the target stimulus according to their mapping structure when the target stimulus was presented. Following the response, participants received feedback for their trial accuracy: if he or she made the correct response, a fixation cross was presented (“inter-trial interval”, ITI) for 500 ms; if an incorrect response was made, the participant saw the mapping structure. Participants advanced the incorrect response screen at their own pace. In each block, participants completed 48 trials. At the end of each block, participants received accuracy (percent correct) and RT (milliseconds) feedback for their overall performance on that block.[1]

Results

Analyses included only data from the second experimental session to minimize effects of learning. Two participants (1 IM, 1 SM) were excluded from the analyses due to low accuracy on the experimental blocks ($<86\%/<500$ correct trials). Accuracy was 97.5% overall in the remaining subjects, equating to fewer than 2 errors per condition. As accuracy approached ceiling in this design, the following analyses were confined to RT.

We limited our RT analysis to trials in which: (1) the response was correct; (2) RT was greater than 200 ms (to avoid fast guesses) and less than 1500 ms (the length of response window before correct responses were shown onscreen); (3) the previous trial was correct; and (4) the stimulus set and the response hand both simultaneously repeated or switched between trials. We used the latter criterion to avoid confounding set-level preparation differences between groups with switching effects, because partial repeats (in which response hand changed but stimulus category repeated or vice versa) were only possible in the one-task but not the two-task group mapping structure. An additional 2 participants (SM group) were excluded for having fewer than 85% usable trials after the application of criteria 1-3.

The remaining participants’ performance was analyzed using a three-way mixed ANOVA including one between-subjects factor for mapping group (IM or SM) and two within-subjects factors: a three-level factor for cue type (stimulus, response, or neutral); and a two-level factor for trial repetition type (switch or repeat). RTs for Experiment 1 are summarized in Table 1. The ANOVA revealed a main effect of trial repetition type ($F(1,28)=34.351$, $p<.001$, $\eta_p^2=.551$). In general, participants were slower when the stimulus and response sets switched from the previous trial. There was also a main effect of cue type ($F(2,56)=24.660$, $p<.001$, $\eta_p^2=.468$). Participants were generally faster for informative than uninformative cues. Cue type further showed a two-way interaction with mapping group ($F(2,56)=3.492$, $p=.037$, $\eta_p^2=.111$), indicating that the effect of cue differed according to whether participants

---

[1] The datasets generated during the current studies are available from the corresponding author on request.

Table 1 Reaction time means and standard deviations for Experiment 1 by mapping group, repetition type, and cue type

	SM		IM	
	RT (ms)	Std. Dev.	RT (ms)	Std. Dev.
Switch				
Neutral	844.6	212.4	765.4	196.8
Stimulus	802.7	211.8	762.9	197.9
Response	796.4	214.5	716.7	198.7
Repeat				
Neutral	816.4	212.5	738.6	185.1
Stimulus	773.0	216.6	714.0	177.2
Response	770.3	198.7	727.8	184.6

SM segregated mapping, *IM* interleaved mapping, *RT* reaction time, *Std. Dev* standard deviation

learned a segregated or interleaved mapping structure. No other effects in the ANOVA were significant.

To better understand the two-way interaction between cue type and mapping group, we analyzed the magnitude of the benefit afforded by the two types of informative cues versus uninformative cues. To calculate the cue benefit for each participant, we subtracted the average RT for each informative cue type from the average RT for neutral cues within each repetition type. The resulting data were submitted to a 2 × 2 × 2 ANOVA, with one between-subject factor for mapping group and two within-subject factors: a two-level factor for repetition type and a two-level factor for cue information (i.e., benefit of stimulus- or response-informative cues). The results demonstrated a main effect of mapping group ($F(1,28)=7.321$, $p=.011$, $\eta_p^2=.207$), indicating that participants in the SM group benefited from informative cues more than participants in the IM group. This was true whether the cues indicated stimulus or response information. This effect is illustrated in Fig. 1.

Perhaps surprisingly, the results in both experiments show a cost of switching between sets which was similar across both groups. Notably, this effect was found in trials in which exact stimulus and response repeats had been removed, which ruled out priming effects. To better understand the role of task representation in the switch costs seen for the IM group, we conducted an additional analysis to investigate how the switch costs might depend on high-level task structure versus individual stimulus and response features. Because the IM group mapping interleaved both stimulus sets across both hands, it was possible to assess how stimulus set and response hand switches separately influenced performance for these participants. We reanalyzed the data, this time including partial switch trials, using a 3 × 2 × 2 repeated measures ANOVA on the RTs from the IM group with factors for cue type (stimulus, response, noninformative), stimulus set repetition, and response hand repetition.

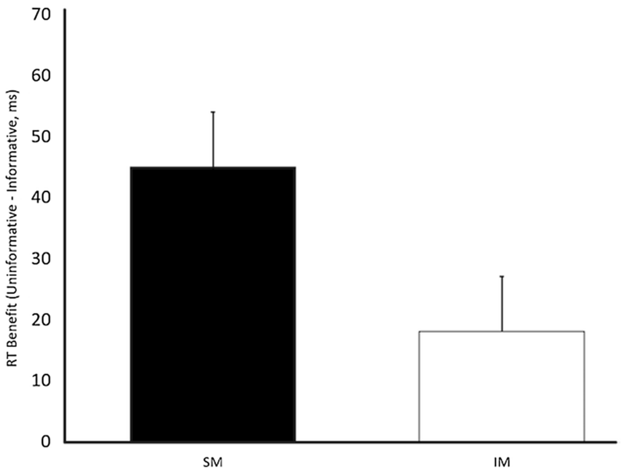

Fig. 1 Cue benefit by mapping group in Experiment 1. Error bars represent standard error of the effect of cue. $F(1,28)=7.321$, $p=.011$, $\eta_p^2=.207$

Table 2 Paired *t* tests between cue types within IM mapping group in experiment 1

Pair	Mean difference (ms)	Std. Dev.	Student's *t*	*p* (2-tailed)
Neutral–stimulus	7.39	14.98	1.972	.067
Neutral–response	21.55	26.29	3.279	.005
Stimulus–response	14.16	30.95	1.831	.087

IM interleaved mapping, *ms* milliseconds, *Std. Dev* standard deviation

First, we found a significant main effect of cue type ($F(1.541,23.114)=6.307$, $p=.010$, $\eta_p^2=.296$). Paired T-tests indicated that this effect was driven by response hand cues (Table 2). These results suggest that cues directly indicating a motor effector allowed for some preparation of that response set, regardless of the other relevant task features. There was also a main effect of response hand repetition ($F(1,15)=17.446$, $p=.001$, $\eta_p^2=.538$). This suggests that there was an inherent cost of switching between hands, regardless of other task factors. Notably, this effect did not interact with cue type ($F(2,30)=1.459$, $p=$n.s.). This means that the cost of switching hands did not depend on the selection of a set at the start of the trial. Furthermore, no other effects were significant; thus, response hand switch costs were not tied to stimulus set switch costs, which would be expected if these costs were dependent on set-level selection processes.

Discussion

In Experiment 1, we manipulated the mapping between identical S–R sets between participants so that half of participants learned a hierarchically structured mapping (i.e., the SM group), while the other half learned an ambiguous mapping that did not afford such structure (i.e., the IM group). Importantly, these mappings were all learned in the context of a single set of task instructions, which did not reference the different structures built into the mappings. We hypothesized that these differences in task structure would lead participants in the SM group to represent the task using a hierarchical task file with separate subgroups for each task, while participants in the IM group would represent it as a flat S–R mapping task.

Because the cue could be used for rule-based selection for the SM but not the IM groups, we predicted that the SM participants would benefit from informative versus uninformative cues more than the IM group. The interaction between cue type and mapping group supports this prediction; the SM group showed a greater benefit of informative cues versus uninformative cues than the IM group. Because both groups performed a task in which they selected a response to a target stimulus from one of eight possible mappings, this difference in cue benefits between groups suggests that the groups represent and implement the task in different ways despite the overlap in stimulus sets, response sets, and task instructions. In other words, the structural organization of the sets of stimuli and responses and their ability to be mapped to distinct high-level, cue-related rules led to a difference in performance in an otherwise equivalent task.

Given these strong cue benefit differences, the equivalent switch costs in this experiment were somewhat surprising. Rogers and Monsell (1995) previously suggested that the switch cost was indicative of the time required to reconfigure processes for a new task. In this case, we might expect switch costs to depend on mapping group in these experiments, as we would expect the SM but not the IM group to show an effect of configuring different subtasks when the stimulus and response sets changed between trials. If this is the case, the equivalent switch costs here would suggest some level of hierarchical set representation for both groups despite their differences in mapping structure. On the other hand, it is possible that these effects are the result of multiple contributing factors, not all of which depend on a hierarchical task structure. Our follow-up analyses within the IM group results suggest that the switch costs in this group do not arise from a hierarchical task representation. In that analysis, the repetition effects were driven by response hand switches, and the small cue benefit (which was, as previously discussed, significantly smaller than that of the SM group) was likewise driven by response hand cues. In fact, stimulus set repetitions/switches did not influence RT alone, nor in combination with other task factors. This suggests that, while switch costs index changes in the factors of a task that must be processed on each trial, they do not necessarily index control processes that are occurring at higher levels of the task set. This is in contrast to the cue benefit, which seems specifically sensitive to set selection processes.

Interestingly, while the SM group demonstrated a larger benefit for informative cues, their RTs were not statistically different on average from the IM group, and in fact were numerically larger across conditions, even for informative cues. At face value, these results appear to violate the HHL; as we have proposed that the IM group should have prepared a set of eight stimulus–response pairs to informative cues while the SM group prepared a subset of four, the HHL would predict that the SM group should have had smaller RTs in the informative condition. However, this prediction assumes that the SM and IM groups had otherwise identical task processing. On the contrary, the task file hypothesis predicts that the SM, but not the IM, group was implementing an additional process to select between two subtasks in addition to, and prior to, the processing required to select the specific response within that set. Participants may then have to operate using this hierarchical structure whether the cue is informative or not; in that case, on uninformative cue trials, participants would nonetheless have to complete the high-level set selection process upon target presentation before they are able to select their specific response. Previous research from our group (Schumacher et al. 2018) has demonstrated a global cost of performing two tasks versus one, which is separate from the trial-by-trial cost of actually switching between those tasks (Mayr 2001). This global cost may index the time required to resolve this additional high-level selection process. The simple act of representing a task as hierarchical subsets rather than as a single flat task may have resulted in slower RTs in the SM group across the task due to the additional demands of resolving the response dictated by a hierarchical task structure. Thus, it is not possible to directly compare RTs between groups.

Given the differences between the groups' performance, we propose that participants in the SM but not the IM group represented the task using a two-level task file, in which cue information selected between two S–R subsets. When participants saw an informative cue, they were able to use the information to select one subset as a whole a priori; this process, then, would not contribute to the RT. This suggests that control processes that produce a cueing benefit occur at the level of high-level rules established during learning of the experimental procedure, even as distinct processes may influence the response level of the task independent of this structure.

Experiment 2

Experiment 1 demonstrated that changing the mapping structure of a task influences participants' ability to use cue information to aid performance. The results suggest that when participants can represent the experimental task hierarchically with subsets of S–R pairs, these sets can be selected between with a high-level rule (as in the SM group) when given a partially informative cue. Participants cannot, however, use these cues when the experimental task does not foster a hierarchical task file structure, as in the IM group. In Experiment 1, the stimulus types were categorically distinguishable and visually complex (viz., faces and buildings). To replicate our findings and investigate whether these cueing benefits occur with less complex and more easily distinguishable stimuli, we repeated the design from Experiment 1 using stimulus sets distinguishable by low-level visual features (viz., location and color). To do this, we replaced the face and place images from Experiment 1 with colored circles and spatial location indicators. Processing of these latter types of stimuli diverges very early in the visual stream to form the ventral and dorsal pathways, respectively (Goodale et al. 1991), whereas the categorical sets from Experiment 1 are both processed within the same ventral pathway (Ungerleider and Haxby 1994).

Methods

Participants

Experiment 2 included 39 participants recruited from the student population at the Georgia Institute of Technology in the same manner as Experiment 1.

Apparatus and stimuli

Apparatus used in this experiment were the same as in Experiment 1. Stimulus images included color and spatial types. The color stimuli consisted of 99 pixel-diameter circles presented in red, yellow, blue, and green. The spatial stimuli consisted of a 100 × 100 pixel crosshair with an asterisk ('*') positioned in one of the four resulting quadrants, 29 pixels from the center of the crosshair in both the x and y directions. All other stimuli in Experiment 2 were the same as in Experiment 1.

Procedure

All experimental procedures proceeded in the same manner as Experiment 1, with the face and place stimulus sets from Experiment 1 replaced with the color and spatial stimulus sets. 21 Participants were assigned an interleaved mapping, and 18 participants a separated mapping.

Table 3 Reaction time means and standard deviations for Experiment 2 by mapping group, repetition type, and cue type

	Segregated		Interleaved	
	RT (ms)	Std. Dev.	RT (ms)	Std. Dev.
Switch				
Neutral	781.7	221.7	813.5	244.8
Stimulus	711.1	221.7	786.1	241.3
Response	713.3	229.4	778.3	252.0
Repeat				
Neutral	713.1	214.7	756.5	237.2
Stimulus	660.1	195.1	736.6	221.5
Response	692.3	111.6	715.6	240.5

SM segregated mapping, *IM* interleaved mapping, *RT* reaction time, *Std. Dev* standard deviation

Results

The same exclusion criteria were applied to these data as in Experiment 1; one participant (IM group) was excluded for accuracy, and an additional 5 (2 in SM group, 3 in IM group) for RT exclusion criteria. Accuracy in Experiment 2 was 95.4% overall in the remaining subjects, equating to fewer than 2 errors per condition; thus, we again limited subsequent analyses to RT.

RT was again analyzed using a three-way mixed ANOVA with factors for mapping group, cue type, and trial repetition type. RTs for Experiment 2 are summarized in Table 3. The ANOVA revealed a main effect of trial repetition type ($F(1,29) = 59.815$, $p < .001$, $\eta_p^2 = .673$). As in Experiment 1, participants performed slower when the trial set switched versus when it repeated. There was also a main effect of cue type ($F(1.818,52.733) = 29.933$, $p < .001$, $\eta_p^2 = .508$). Once again, participants were faster for informative than uninformative cues. Critically, cue type also showed the same two-way interaction with mapping group ($F(1.818,52.733) = 5.422$, $p = .009$, $\eta_p^2 = .158$) as in Experiment 1. No other effects in the ANOVA were significant.

We again conducted a second analysis to assess the difference in cue benefits as a function of cue information using a three-way mixed ANOVA with factors for mapping group, repetition type, and cue information (stimulus cue benefit vs response cue benefit). The results again demonstrated a main effect of mapping group ($F(1,29) = 7.410$, $p = .011$, $\eta_p^2 = .204$), indicating that participants used cues differently, regardless of what information they carried, depending on their mapping structure. This effect is

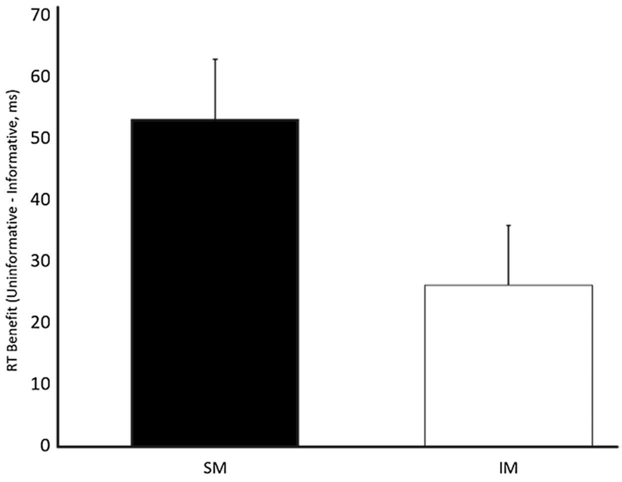

Fig. 2 Cue benefit by mapping group for Experiment 2. Error bars represent standard error of the effect of cue. $F(1,29)=7.410$, $p=.011$, $\eta_p^2=.204$

Table 4 Paired *t* tests between cue types within IM mapping group in Experiment 2

Pair	Mean difference (ms)	Std. Dev.	Student's *t*	*p* (2-tailed)
Neutral–stimulus	21.02	21.08	3.987	.001
Neutral–response	30.70	41.86	2.934	.010
Stimulus–response	9.69	11.89	.815	.428

IM interleaved mapping, *ms* milliseconds, *Std. Dev.* standard deviation

illustrated in Fig. 2. In addition, there was a main effect of repetition type ($F(1,29)=4.980$, $p=.034$, $\eta_p^2=.147$), which indicated that the cue benefit was larger for switch trials than for repeat trials (not shown). No other effects were significant.

Finally, we conducted the follow-up analysis within the IM group using three-way repeated-measures ANOVA with factors for cue type; stimulus set repetition type, and response hand repetition type. Again we found a main effect of cue type ($F(1.524,22.865)=6.438$, $p=.010$, $\eta_p^2=.300$); in this case, both stimulus and response cues showed faster RTs than noninformative cues (Table 4). There were also significant main effects of both stimulus ($F(1,15)=37.217$, $p<.001$, $\eta_p^2=.713$) and response set repetition type ($F(1,15)=11.009$, $p=.005$, $\eta_p^2=.423$), as well as a two-way interaction between these two factors ($F(1,15)=16.027$, $p=.001$, $\eta_p^2=.517$). There were small independent costs of switching stimulus set or response hand; the largest effect was a benefit of repeating both sets simultaneously. No other effects were significant.

Discussion

The results in Experiment 2 partially replicate those in Experiment 1. As in Experiment 1, RT was influenced by both cue type and trial switches, such that RTs were faster for informative cue types and for trial repeats. Importantly, only cue type interacted with mapping group, indicating that the SM but not the IM group demonstrated cue benefits consistent with a hierarchical selection process. Thus, despite the differences in the features of the stimulus sets, performance in both tasks depended on the task structure afforded by the S–R mapping. Taken together, the results from Experiments 1 and 2 suggest that the organization of the responses for the S–R pairs affects the representation of a task and influences the ability to use a partially informative cue—even for very distinguishable stimuli.

Additionally, both groups again showed similar switch costs. However, follow-up analyses of performance in the IM group indicated a different source of these costs in Experiment 2. Whereas switch costs in the IM group in Experiment 1 were driven by response hand features, in Experiment 2 they depended jointly on both the stimulus and response sets, such that performance suffered to a similar degree if any dimension of the target or response was changed. This could suggest that the IM group was able to form coherent set representations of S–R groups. However, the persistent difference in cue benefits between groups nonetheless indicates that such a set-based representation is insufficient to support the rule-based set selection processes afforded by a hierarchical task representation.

Notably, we did not see the same overall slowing in the SM group versus the IM group in this experiment that was present in Experiment 1, and in fact the RTs in the informative cue condition showed a between-group pattern consistent with the predictions of the HHL. This suggests that the slowing in Experiment 1 was partially due to the complexity of our face and scene stimuli. Recent research in the visual working memory field has shown that complex objects in working memory are represented by content-free labels, or pointers (in fact, similar to the concept of a task file), which guide online retrieval of the specific features of the object when the full representation is required (Huang and Awh 2018). This process may interact with task set representations to produce the exacerbated global switch costs observed in Experiment 1. Specifically, in tasks with complex stimuli, the "stimulus" in a stimulus–response mapping may in fact be one of these content-free labels, which would then have to be mapped back onto the individual target features for each stimulus. Global switch costs are thought to be due to the additional control demands of representing multiple sets, as discussed in Experiment 1; when these sets use complex stimuli, then, they would have even greater representational demands, resulting in a super additive cost in the SM group over the IM group. In Experiment 2, on the other hand, the stimulus features could be more directly bound with their respective responses, minimizing this additional processing requirement.

General discussion

The present experiments investigated how the structure of a task influences the implementation of control processes for response preparation. Across two experiments, sets of stimuli and responses were associated with a mapping structure that encouraged the representation of the task as two hierarchically linked subsets (the SM group) or as a single overarching task (the IM group). We found that, regardless of the stimulus type, the SM group benefited more from partially informative cues more than the IM group. Specifically, the SM group showed a higher cue benefit than the IM group, which would violate the HHL if cue information could be used to reduce the number of S–R pairs regardless of task structure. Instead, cues appear to be more readily utilized when the task mapping lends itself to a hierarchical representational structure. In other words, participants were best able to use a mnemonic cue that indicated semantic information about either the stimulus category or response hand when that information indicated a distinct subset of the stimuli and responses. The results from these experiments provide converging evidence for the task file hypothesis.

Task file hypothesis versus the grouping model

While the set-based mapping structures described in the task file account are reminiscent of the subgroups of the grouping model (Adam et al. 2003) insofar as both include subsets of task-relevant S–R pairs, the latter hypothesis is limited in its ability to capture the mechanisms underlying cognitive control. The grouping model describes how cue benefits arise for exogenous cues, or those with direct spatial correspondence with the S–R groupings. However, in the present experiments, the cues provide only semantic information about the stimulus category. Therefore, participants in this study must translate the cue stimulus into a task relevant representation before using this information to prepare a response set. The grouping model posits that exogenous cues cause participants to shift their attention to one S–R subgroup via “attentional zooming”; in other words, because the cues appear in the positions of two potential stimuli that compose a subgroup of the task, participants limit their upcoming selection process to this pre-activated subset. However, this does not account for the additional translational step required to interpret the semantic information of a mnemonic cue to use it for subtask selection in the current experiments.

On the other hand, the task file hypothesis accounts for benefits of both exogenous and mnemonic cues parsimoniously. In our design, participants in the SM group represent the task as two subtasks, which they then can select between at a symbolic level when presented with cue (viz. Schumacher and Hazeltine 2016; Hazeltine and Schumacher 2016), while participants in the IM group represent the task with a flat structure and cannot use the cue to select a distinct subgroup of S–R pairs. This difference in ability to define the selection rule for choosing between subsets of a task leads to differences in the ability to benefit from an informative cue. Thus, the task file boundaries define the implementation of control.

Role of task structure in cognitive control

The experimental discussions above have suggested that switch costs in this task may index cognitive control processes that are distinct from those indexed by the cue benefit. More specifically, it seems that cue benefits reflect control processes that occur at the set level to support rule-based set selection, while switch costs are the result of control processes that are invoked when one switches between sets with different features. This distinction between task-level and feature-level influences on cognitive control is well-supported by the literature.

Hazeltine, Lightman, Schwarb, and Schumacher (2011) investigated set-level effects using the congruency sequence effect (CSE, referred to in that report as “sequential modulations”); the CSE should only transfer within a task set, but not between them (Akçay and Hazeltine 2007). They found that CSE transfer was modulated by presentation modality in some designs, indicating that participants could use modality to form a set-level task boundary. Critically, the CSE was only bounded by modality set in experiments in which responses were uniquely mapped to a single stimulus identity. That is, participants were able to form modality-based sets only when they could prepare a subset of nonoverlapping stimulus–response mappings based on the modality information provided by the distractor (which in the prime-probe task also serves as a cue). On the other hand, item-level effects, as measured by the influence of flanker history on performance, occurred in all experiments and were not bounded by modality regardless of the design. These results suggest that cognitive control is not a monolithic process, but a conglomeration of processes that affect performance on a task by influencing different levels of the hierarchical representation.

Understanding what determines task structure

Previous research by our group has suggested a mechanism by which task sets are instantiated in the brain that suggests a rationale for predicting whether or not a task will be represented hierarchically (Cookson et al. 2016). In that experiment, participants were given the SM mapping structure described in Experiment 1. Using an event-related design that isolated cue-related activity, we found brain activation in task-relevant response-processing regions, as well as increased connectivity between these areas and task-relevant stimulus processing regions, as a function of the subtask indicated by the cue (e.g., fusiform face area and premotor cortex). That is, when participants saw a cue for, e.g., an upcoming face stimulus, their fusiform face area increased connectivity with the motor region associated with the upcoming response hand and upregulated activity in the latter area. These changes did not occur in and between regions that were task relevant, but were not indicated by the cue as relevant for the current trial.

These results suggested that task sets were instantiated on a trial by trial basis by positively modulating the connectivity within a subnetwork of regions that were relevant to the task. Importantly, the task sets in that experiment were associated with two pairs of distinct stimulus- and response-processing regions that did not overlap. This ostensibly minimized the amount of remapping required to switch between sets. Task boundaries, then, may depend on the ability to map subtasks to distinct networks in the brain. In the IM group, this would not be possible; networks connecting to each response hand region would necessarily overlap in both stimulus processing regions to support the two stimuli interleaved from each set.

Notably, unlike in the experiments reported here, the set-level effects from Hazeltine et al. (2011) did not require distinct response sets for participants to establish distinct task sets. That is, the visual and auditory versions of the stimuli from those experiments were mapped to the same response sets. At first glance, this may seem at odds with the above explanation; however, there is a key difference between these designs. In that design, the auditory and visual versions of each stimulus nonetheless represented the same stimulus identity. As described in the Experiment 2 discussion, perceptual objects can be represented using content-free labels, which are stored independent of their perceptual features (Huang and Awh 2018). If this is the case, then the stimulus–response mappings in Hazeltine and colleagues' design could have been represented in a network that statically linked responses to these content-free labels, regardless of the input modality. On any given trial, then, participants simply needed to connect this network to either the visual or the auditory processing area as a function of the modality information contained in the distractor.

Conclusion

The task file hypothesis provides an elegant and parsimonious explanation of the complex behavioral phenomena surrounding cognitive control. Representing a task using a hierarchical structure allows for the highly flexible coordination of different actions according to current context and ongoing goals. The current experiment demonstrates participants' propensity to represent tasks using these hierarchical structures automatically where the task design allows, even when other task features are equal, and that this difference in representation has meaningful impacts on behavioral measures. Furthermore, these structure-based impacts are distinct from feature-related impacts, suggesting that different hierarchical levels of a task are affected by distinct aspects of cognitive control processes.

References

Adam, J. J., Hommel, B., & Umiltà, C. (2003). Preparing for perception and action (I): The role of grouping in the response-cuing paradigm. *Cognitive Psychology*. https://doi.org/10.1016/s0010-0285(02)00516-9.

Adam, J. J., Hommel, B., & Umiltà, C. (2005). Preparing for perception and action (II): Automatic and effortful processes in response cueing. *Visual Cognition*. https://doi.org/10.1080/13506280444000779.

Akçay, Ç., & Hazeltine, E. (2007). Conflict monitoring and feature overlap: Two sources of sequential modulations. *Psychonomic Bulletin & Review, 14*(4), 742–748. https://doi.org/10.3758/bf03196831.

Cookson, S. L., Hazeltine, E., & Schumacher, E. H. (2016). Neural representation of stimulus–response associations during task preparation. *Brain Research, 1648*(Part A), 496–505. https://doi.org/10.1016/j.brainres.2016.08.014.

Fitts, P. M., & Seeger, C. M. (1953). S–R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*(3), 199–210. https://doi.org/10.1037/h0062827.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*(6305), 154–156. https://doi.org/10.1038/349154a0.

Goodman, D., & Kelso, J. S. (1980). Are movements prepared in parts? Not under compatible (naturalized) conditions. *Journal of Experimental Psychology: General, 109*(4), 475–495. https://doi.org/10.1037/0096-3445.109.4.475.

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1898.

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.), *The psychology of learning and motivation* (Vol. 64, pp. 195–245). San Diego, CA, US: Elsevier Academic Press.

Hick, W. E. (1952). On the rate of gain of information. *The Quarterly Journal of Experimental Psychology, 4,* 11–26. https://doi.org/10.1080/17470215208416600.

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports, 8*(1), 23. https://doi.org/10.1038/s41598-017-18157-5.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*(3), 188–196. https://doi.org/10.1037/h0056940.

Martinez, A. M., & Benavente, R. (1998). The AR face database. *CVC Technical Report, 24*.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging, 16*(1), 96.

Miller, J. (1982). Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 273–296.

Miller, J. (1983). Can response preparation begin before stimulus recognition finishes? *Journal of Experimental Psychology: Human Perception and Performance, 9*(2), 161–182.

Proctor, R. W., & Reeve, T. G. (1985). Compatibility effects in the assignment of symbolic stimuli to discrete finger responses. *Journal of Experimental Psychology: Human Perception and Performance, 11*(5), 623–639. https://doi.org/10.1037/0096-1523.11.5.623.

Reeve, T., & Proctor, R. (1984). On the advance preparation of discrete finger responses. *Journal of Experimental Psychology: General, 10*(4), 541–553.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109*(4), 444.

Rosenbaum, D. A. (1983). The movement precuing technique: Assumptions, applications, and extensions. *Advances in Psychology, 12,* 231–274.

Schumacher, E. H., Cookson, S. L., Smith, D. M., Nguyen, T. V. N., Sultan, Z., Reuben, K. E., et al. (2018). Dual-task processing with identical stimulus and response sets: Assessing the importance of task representation in dual-task interference. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.01031.

Schumacher, E. H., & Hazeltine, E. (2016). Hierarchical Task Representation: Task Files and Response Selection. *Current Directions in Psychological Science*, *25*(6), 449–454. https://doi.org/10.1177/0963721416665085.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology, 4*(2), 157–165.