Cognition 182 (2019) 251–259

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

How metacontrol biases and adaptivity impact performance in cognitive search tasks☆

Vera N. Mekern*, Zsuzsika Sjoerds, Bernhard Hommel

Leiden University, Institute of Psychology, Cognitive Psychology Unit, & Leiden Institute for Brain & Cognition, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands

ARTICLE INFO

Keywords:
Cognitive control
Metacontrol
Response production tasks
Persistence
Flexibility
Adaptivity

ABSTRACT

Cognitive control requires a balance between persistence and flexibility. We studied inter- and intraindividual differences in the metacontrol bias towards persistence or flexibility in cognitive search tasks from various cognitive domains that require continuous switching between persistence and flexibility. For each task, clustering and switching scores were derived to assess persistence and flexibility, respectively, as well as a total performance score to reflect general performance. We compared two, not mutually exclusive accounts according to which the balance between clustering and switching scores is affected by (1) individual, trait-like metacontrol biases towards persistence or flexibility and/or (2) the metacontrol adaptivity to bias states according to changing situational demands. We found that clustering and switching scores failed to generalize across tasks. However, clustering and switching were inversely related and predicted the total performance scores in most of the tasks, which in turn partially generalized across tasks and task domains. We conclude that metacontrol-biases towards persistence or flexibility can be adapted easily to specific task demands and individual resources, possibly overwriting individual metacontrol trait biases. Moreover, we suggest that total performance scores might serve to measure metacontrol adaptivity in future studies if task-restrictions and resources are known and/or well balanced.

1. Introduction

Intentional agents would be ill-advised to rely too much on cognitive persistence (needed to reach longer-term goals); they also need some degree of flexibility to register, and switch to, alternative opportunities (Goschke, 2003). For example, while increased top-down control benefits performance on tasks that require cognitive persistence (e.g., Fischer & Hommel, 2012) performance on tasks that require more cognitive flexibility might suffer from increased top-down control (e.g., Bocanegra & Hommel, 2014; Stock, Steenbergen, Colzato & Beste, 2016; Zink, Stock, Colzato, & Beste, 2018). Research provides evidence for intraindividual variability in cognitive control, as humans can adjust their bias towards persistence or flexibility (Dreisbach & Goschke, 2004; Hommel & Colzato, 2017; Zink et al., 2018). Various factors that promote particular biases have been identified, including mood (Dreisbach & Goschke, 2004), meditation-induced states (Colzato, Ozturk, & Hommel, 2012; Colzato, Sellaro, Samara, & Hommel, 2015; Colzato, Szapora, Lippelt, & Hommel, 2017; Colzato, van der Wel, Sellaro, & Hommel, 2016), and reward (Hefer & Dreisbach, 2016, 2017). Neuroscientific evidence suggests that such metacontrol biases (to use the terminology suggested by Hommel & Colzato, 2017) are regulated through the interplay of frontal and striatal dopaminergic pathways (Cools & D'Esposito, 2011; Cools, 2015) and/or the interplay between dopaminergic D1-receptors and D2-receptors (Durstewitz & Seamans, 2008).

Humans do not only show intraindividual variability in their metacontrol *state bias*, but they also differ systematically in their individual metacontrol default or *trait bias*: Differences in both genetic setup (in genes relevant for dopaminergic processing) and cultural background have shown to be associated with particular biases towards persistence or flexibility (Hommel, Colzato, Scorolli, Borghi, & van den Wildenberg, 2011; for a review, see Hommel & Colzato, 2017).

The persistence/flexibility tradeoff is thus likely to emerge from some interplay between inter- and intraindividual (i.e., trait and state)

Abbreviations: WPT, word production task; AUT, Alternative Uses Task; 5PT, five point task; VST, verbal search task; MAB, multi-armed bandit task; RAT, Remote Associates Test

☆ This work was supported by the European Research Council (ERC Advanced Grant: grant agreement No. 694722 [Metacontrol] to B. Hommel) under the European Union's Horizon 2020 research and innovation program.

* Corresponding author at: Leiden University, Institute of Psychology, Cognitive Psychology Unit, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands.
E-mail addresses: v.n.mekern@fsw.leidenuniv.nl (V.N. Mekern), z.sjoerds@fsw.leidenuniv.nl (Z. Sjoerds), hommel@fsw.leidenuniv.nl (B. Hommel).

https://doi.org/10.1016/j.cognition.2018.10.001
Received 29 May 2018; Received in revised form 30 August 2018; Accepted 3 October 2018
0010-0277/ © 2018 Elsevier B.V. All rights reserved.

metacontrol biases, and it was this interplay that we aimed to investigate in the present study. We tracked inter- and intraindividual differences across a range of tasks that arguably require the continuous adjustment of metacontrol biases. Particularly, we considered six tasks that require a cognitive search in which the persistence/flexibility tradeoff should be made. Three of these tasks were fluent response production tasks (often termed *fluency tasks*) which are fairly unrestricted in terms of external stimuli and instructions. Participants produced words (word production task; WPT; Troyer, Moscovitch, & Winocur, 1997), designs (five point task; 5PT; Regard, Strauss, & Knapp, 1982), or ideas (e.g., uses for an everyday object; Alternative Uses Task; AUT; Guilford, 1967) in response to one single cue stimulus. While general performance on these fluent production tasks can be measured in terms of a *total performance score* (the total amount of responses), the balance between persistence and flexibility in these tasks might be reflected by measures of clustering and switching, respectively. For example, in the phonemic version of the WPT people tend to respond in clusters of words that are phonemically similar (e.g., *fact*, *factor*, and *face*) and participants with a metacontrol bias towards persistence might be inclined to cluster more than those with a bias more towards flexibility (e.g., Troyer et al., 1997).

The three other tasks were arguably more restricted by instructions or stimuli. Despite terminological differences, all these tasks have been used to assess aspects of cognitive persistence and flexibility. We included the verbal search task (VST) introduced by Hills, Todd, and Goldstone (2008, 2010), used to measure the exploration/exploitation tradeoff in cognitive search, but in which pre-defined letter sets restrict the patch of words participants can search from. We included a multi-armed bandit task (MAB) thought to assess the exploration/exploitation tradeoff (e.g., Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Jepma, Beek, Wagenmakers, van Gerven, & Nieuwenhuis, 2010; Jepma & Nieuwenhuis, 2011). In this task, individuals engage in exploitation of high-value slot machines or in exploration between four machines when payoff changes over time. Finally, we included the Remote Associates Test (RAT; Mednick, 1962), which was designed to focus more on convergent thinking, and thus qualifies as a rather persistence-heavy search task (e.g., Colzato et al., 2017; Fischer & Hommel, 2012; Guilford, 1950, 1967). However, performance on this task benefits from switching between problems, indicating that a persistence/flexibility tradeoff should be realized (Lu, Akinola, & Mason, 2017).

Considering the probable interplay between inter- and intraindividual differences in controlling the persistence/flexibility tradeoff, one could think of two different, but not necessarily mutually exclusive, ways that individual differences in metacontrol bias express themselves in cross-task correlations. One possibility is that performance in cognitive search tasks is strongly affected by an individual trait bias: Individuals with a strong persistence bias would be more likely than others to exploit or cluster, while individuals with a strong flexibility bias would be more likely than others to explore or switch (Fig. 1). Indeed, clustering and switching in verbal search has been reported to be related to clustering and switching behavior in a visual search task (Hills et al., 2008, 2010), suggesting that these tasks share enough similarity to be sensitive to possible trait biases. To reveal such possible trait biases, we analyzed indicators of persistence (clustering, exploitation) and of flexibility (switching, exploration) separately, so that possible cross-task correlations could indicate trait biases towards persistence or flexibility.

Another possibility is that cross-task correlations are limited to total performance scores, possibly reflecting the degree to which people adapt their metacontrol state bias to match situational demands. For instance, in the WPT, if a participant has a large vocabulary, a good strategy could be to mostly exploit/cluster, as this person can produce more words within one semantic or phonemic cluster, and thus has to explore/switch to other clusters of words only occasionally (Fig. 2, panel A; Unsworth, Spillers, & Brewer, 2011). In contrast, for a participant with a small vocabulary, the opposite strategy could be more suitable. If so, we might find significant cross-task correlations for clustering and switching for vocabulary-dependent tasks at most, but no generalization beyond those, suggesting no strong evidence that the balance between clustering and switching solely depends on the metacontrol trait bias across different types of tasks. However, individuals might differ with respect to the degree to which they take their resource limitations (such as the size of their vocabulary) into account and adapt their metacontrol bias to match task demands (Fig. 2, panel B). We suggest that this ability, which we will refer to as *adaptivity*, would be reflected in generalizability of total performance rather than of the clustering/exploitation and switching/exploration score, as total performance might depend on how much individuals adapt the balance between clustering and switching accordingly to match the situational demands. Interestingly, generalizability over fluent response production tasks in different domains of cognitive functioning has been studied mainly in terms of total performance scores which have been found to generalize over tasks (e.g., Ardila, Rosselli, & Bateman, 1994; Unsworth et al., 2011; Vannorsdall, Maroof, Gordon, & Schretlen, 2012; Whiteside et al., 2016; but see Schmidt et al., 2017). This suggests that there is intraindividual stability in overall performance regardless of domain, or, according to our reasoning, an indication of the degree of metacontrol bias adaptivity. To test whether interindividual differences related to metacontrol reflect the individual degree of adaptivity, rather than (or in addition to) stable interindividual differences in trait biases, we thus analyzed not only clustering and switching scores but also total performance.

Related to the idea that clustering and switching are adaptively balanced to increase total performance, we also tested to which degree the total performance scores are predicted by clustering and switching scores. We additionally explored whether two non-invasive alleged proxy-measures of individual dopamine levels would be statistically related to clustering, switching, and total performance scores, which we report in the Supplementary material.

To summarize, according to the trait-bias account, one would expect that clustering and switching scores would correlate across tasks in different domains. However, according to the adaptivity account, one

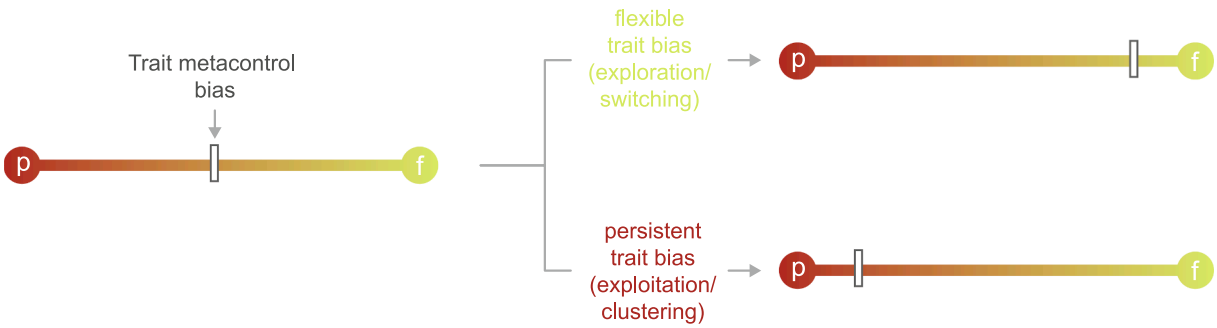

Fig. 1. Differences in genetic setup and cultural influences might induce trait biases in metacontrol. Those individuals with a flexible metacontrol bias should be more likely to explore alternatives and switch between them while those with a persistent bias should be more prone to exploit and cluster.

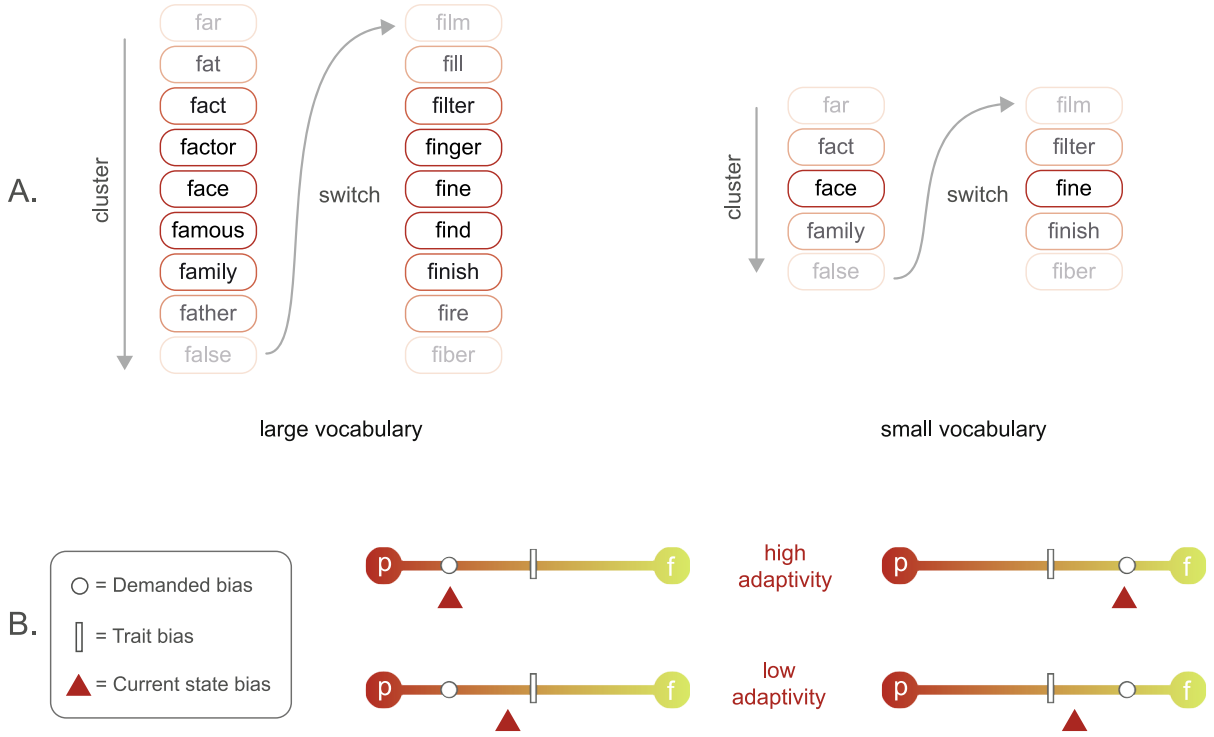

Fig. 2. A. The same task can require different individuals to adapt their metacontrol state to different metacontrol biases. In the word production task, individuals with a large vocabulary (left) might switch less between phonemic clusters of words, as they are able to produce more words within one cluster, suggesting a persistence bias, while for individuals with a small vocabulary (right) the opposite might be true. B. Moreover, differences in the way individuals adapt to situational demands (considering task demands and individual resources) can lead some individuals to change the balance between clustering/switching from task to task (potentially high adaptivity) while others stay near their default balance, dictated by their metacontrol trait bias (potentially low adaptivity).

would expect that total performance scores would correlate across tasks and domains but clustering and switching scores correlate only across tasks within domains or not at all.

2. Methods

2.1. Participants

We collected data from 160 participants at the Leiden University Institute of Psychology. Of the tested sample, 31 participants were excluded before analysis. Post-hoc exclusion reasons were recent drugs use (e.g., dopaminergic, cholinergic, anxiolytic, or illicit drugs; $n = 12$), a history of psychiatric disorders ($n = 8$), dyslexia ($n = 6$), non-nativity ($n = 1$), serious color blindness ($n = 1$), and other neurological issues (e.g., unspecified neurological disorders or serious concussions; $n = 3$). The final sample ($N = 129$, 87 females, $M_{age} = 21.43$, $SD_{age} = 2.37$) consisted of native Dutch speakers in self-reported excellent mental and physical health. Due to additional missing data because of mostly technical problems and the fact that only half of all participants performed the multi-armed bandit task and the verbal search task, sample sizes vary over analyses (Supplementary material). All participants signed informed consent for the study, which was approved by the local psychology research ethics committee (Leiden University, Institute of Psychology). Participants received study credit or a monetary reward (13 EUR).

2.2. Instruments

2.2.1. Word production task

The word production task (WPT) required participants to respond with as many words as they could think of, starting with *L*, *B*, or *S* in one minute per letter (also known as *verbal fluency*; Troyer et al., 1997). Consecutive words were considered part of a cluster when they started with the same two letters, rhymed, differed only by a vowel sound, or were homonyms (Troyer et al., 1997). Participants were asked not to respond with proper words or variants of the same word (such as *fat* and *fatter*).

2.2.2. Alternative Uses Task

In the AUT (Guilford, 1967) participants named as many uses for everyday objects (*pen* and *towel*) as they could think of within five minutes. Consecutive uses were considered part of a cluster when they were related according to shape and/or a specific use, e.g., a cluster of uses of a pen could be *a miniature lighthouse* and *a miniature lamp post* (Gilhooly, Fioratou, Anthony, & Wynn, 2007).

2.2.3. Five-point design production task

In the five-point task (5PT), participants created as many designs as possible in two minutes by connecting five dots by single lines (also known as *design fluency*; Fig. 3; Regard et al., 1982; Tucha, Aschenbrenner, Koerts, & Lange, 2012). Consecutive designs were considered part of a cluster according to three possible strategies (Gardner, 2008; Vik & Ruff, 1988): a rotational strategy (the design or part of it is rotated), a quantitative strategy (single lines were systematically added or subtracted), and a blended strategy (the strategies combined; Gardner, 2008). Participants were asked not to repeat designs and to use only single, straight lines that connected two dots.

In the WPT, AUT and 5PT the total performance score was the total number of responses. Clustering was the average number of words, designs, or ideas in a cluster, starting to count from the second consecutive response in each cluster relative to the total number of responses in multi-response clusters only. For example, in the word production task, a phonemic cluster in the series *'film, far, fat, fabulous, fork'* would start with *far*, end with *fabulous*, and cluster size would be 2 (based on Troyer et al., 1997). Switching was the number of times participants switched between single responses or clusters of responses.

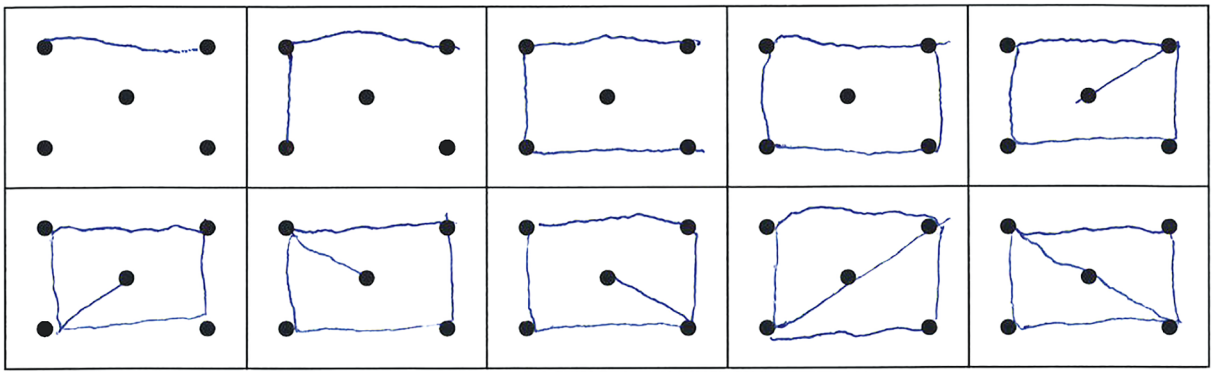

Fig. 3. Ten consecutive designs in the five-point task (Regard et al., 1982). The first eight designs form a cluster according to a blended strategy of a quantitative strategy (first five designs) continuing into a rotational strategy (fifth through eighth design; Gardner, 2008).

Total performance, clustering, and switching scores were standardized in each task. Errors and perseverations were included in each fluent production task as they still reflected the use or the lack of clustering strategies. Moreover, earlier work shows that perseverations are not related to clustering or switching, suggesting that individual differences in errors are not related to clustering and switching either (Unsworth et al., 2011).

2.2.4. Verbal search task

The verbal search task (VST) measures exploration/exploitation strategies in verbal search and was used to show that search processes are generalizable over different modalities (Hills et al., 2008, 2010). Participants were instructed to find Dutch words of at least four letters using only letters from a 6-letter set presented on the computer screen. Each letter could be used once per word and words could not be plurals or proper names. Participants could continue to the next letter set when they felt they had exploited the current set but had to wait 15 s between sets to resemble travel time in exploration. They were told to use as much time as necessary, but to not stick around too long or too short in each set. Participants were shown a maximum of 14 letter sets, presented in random order, from which they were to find a maximum of 30 correct words. Feedback was presented for 800 ms in each trial on whether the word was correct, and the total amount of correct words was displayed continuously. We regarded the number of switches between lettersets as a measure of switching behavior and the average number of words per set was our measure of clustering. We recorded the time participants searched each set and the total time required to find 30 words was the total performance score. Before the VST participants performed a visual search in either a clumpy or diffuse environment. To correct for possible priming influences of this difference the scores in verbal search were all standardized per condition prior to analyses. We multiplied all standardized total performance scores in the VST with -1, such that a higher score indicated better performance, similar to the total performance scores in other tasks.

2.2.5. Multi-armed bandit task

Participants played four slot machines to gain as many points as possible (Daw et al., 2006). The slot machines were displayed in each corner of a computer screen, and participants chose an arm by pressing the Q, W, S, or A. Over 200 trials, mean payoff of each arm gradually changed such that participants should continuously re-adjust their exploration/exploitation strategy to track the highest paying arm (Daw et al., 2006; Jepma & Nieuwenhuis, 2011; Jepma et al., 2010). Reward was displayed for 800 ms in each play, and total reward was displayed continuously. There was no time limit, but on average the task lasted for approximately 5 min. We proposed that participants clustered when they played the same arm for two or more consecutive trials and calculated cluster size as the average size of clusters (starting to count from the second consecutive play) relative to the total number of clusters. Switching was the number of switches between arms, counting only exploratory switches (i.e., in the series of plays *arm1-arm1-arm1-arm2-arm1* the last switch from *arm2* back to *arm1* is not exploratory). As participants played one of two different versions of this task, containing different, but comparable, random walks for pay-out per arm, all scores were standardized per version prior to analyses (Jepma et al., 2010).

2.2.6. Remote Associates Test

A 22-item, Dutch, pen-and-paper version of the RAT was used (Akbari Chermahini, Hickendorff, & Hommel, 2012; Mednick, 1962). This task is mostly used to test convergent thinking, however, recent work shows that switching between items helps to solve them quicker suggesting a necessary tradeoff between persistence and flexibility (Lu et al., 2017). Each item consists of three seemingly unrelated words that the participant is required to connect, by thinking of a fourth word that can be combined with each of the three stimulus words (e.g., in English, the words *dew/comb/bee* should all be combined with the word *honey*). The participant was required to solve as many items as possible within 5 min. The final score was the number of correctly solved items. No clustering or switching scores were available for our pen-and-paper version of the RAT. However, the final score might be insightful still, especially regarding the adaptivity account.

2.3. Analyses

To address our first aim, we analyzed whether clustering and switching measures generalized over tasks and thus positively correlated between the WPT, 5PT, AUT, VST, and the MAB. For our second aim, we analyzed whether total performance scores correlated between all tasks and followed up with an exploratory factor analysis to gain more insight into this correlational structure in a sample with complete scores on all tasks. For our third aim, we tested whether clustering and switching could be interpreted to reflect two ends of a balance between which the balance can be adapted to perform well in terms of total performance. To do so, we statistically predicted total performance from cluster size and switching within each task in a multiple regression analysis to study whether clustering and switching indeed predicted total performance together. We also tested whether clustering and switching within each task were inversely related by finding partial correlations corrected for total performance.

All analyses that address dopamine as a possible neuromodulator of clustering, switching, and total performance, as well as related results can be found in the Supplementary material. For completeness we also report post-hoc calculated power analyses for the correlational and regression analyses in the Supplementary material.

Considering the exploratory nature of our study all hypotheses were tested two-sided and alpha was set at .05 for every hypothesis test. When assumptions for parametric tests were violated we followed up with non-parametric or robust testing. Unless stated otherwise, this did

not change interpretation and the discussed results are parametric test results.

3. Results and discussion

Interrater reliability for cluster size and switching scores in the WPT, 5PT and AUT was based on a random sample of data that was scored by a second rater. Both raters were trained psychologists. According to type 2 intraclass correlation coefficients of consistency (ICC2) reliability in the WPT ranged from good (between .60 and .74; Cicchetti & Sparrow, 1981) to excellent (between .75 and 1.00). Reliability of the switching scores was slightly higher (ranging from .86 for *L* to .97 for *B*) than those for cluster size (ranging from .66 for *L* to .92 for *B*). Interrater reliability in the 5PT was good for both switching (ICC2 = .79) and cluster size (ICC2 = .71). In the AUT, interrater reliability for switching was excellent ($ICC2_{towel}$ = .97 and $ICC2_{pen}$ = .98), and good for cluster size ($ICC2_{towel}$ = .62 and $ICC2_{pen}$ = .85).

3.1. Generalizability of clustering and switching between tasks

In light of our first aim we tested the generalizability of clustering and switching scores by studying zero-order correlations. Clustering scores did not correlate between any two tasks (Table 1), and switching positively correlated only between AUT and 5PT (Holm-corrected for multiple comparisons; $r_{spearman}$ = .344, $CI_{95\%}$ = [.18,.49], $p < .001$; Table 2). This counters expectations based on the trait-bias account, but might support the adaptivity account, as the absence of correlations between clustering as well as switching scores indicates that participants are biased towards clustering or switching to a different extent depending on the task, and hence adapt to the task demands. We should take into account that some pairwise correlations were calculated in smaller samples, which might have led to type-II errors. Nonetheless, if the correlations in these smaller subsamples would have been significant our conclusions would remain the same.

3.2. Generalization of total performance scores over tasks

While the absence of cross-task correlations between clustering and switching scores indicates that these scores do not reflect a metacontrol trait bias, putative support for the adaptivity account would be provided if total performance scores would generalize over tasks, as this might reflect the ability of individuals to adapt their metacontrol state to the specific task demands. Although total performance was not related across all tasks (Table 3), the unrestricted fluent production tasks (WPT, AUT, and 5PT) were significantly and positively correlated to each other. Moreover, performance in the WPT was related to performance in the VST. These correlations were based upon pairwise deletion, such that each coefficient was based on a different sample. Again, in the smaller subsamples we have to be aware of possible type-II errors. However, if correlations in these subsamples would have been significant this would offer more support for our conclusions. We followed up with an exploratory factor analysis in the subsample of participants (n = 63) that had complete data on all tasks. Assumption checks showed sufficient correlational structure to be able to find factors (Bartlett's $\chi^2(15) = 41.439, p < .001$), no sign of multicollinearity (determinant = .459), and adequate sampling (Kaiser-Meyer-Olkin = .72 overall, with values ranging from .64 to .79). Our data was not multivariate normally distributed, therefore we used principal axis factoring.

A parallel analysis, the scree plot, and Kaiser's rule all suggested a two-factor solution (Table 4). Mean communality was low at .35, especially RAT and 5PT showed low communality, suggesting that only a small part of the variance in these tasks is related to the same underlying factors compared to the rest of the tasks. Combined with moderate factor loadings and small sample size (which may have caused a type-II error in the χ^2-test) these results should only be interpreted with care. However, the residual matrix showed that only 6.67% of all off-diagonal residuals was larger than |0.05|. Moreover, we repeated the analysis with a maximum likelihood factoring method to establish that another type of factoring method confirmed the currently discussed findings (Supplementary material). While the two-factor solution ($\chi^2(4) = 1.02$, RMSEA = 0, 90% CI: [0, .069], TLI = 1) indicated that variance in the tasks is caused by two different underlying factors, the correlation between factors was of medium size (r = .413) indicating that there was overlap in variance, such that this finding still offers support for our adaptivity account. However, the fact that the tasks loaded onto two separate factors suggests that these scores of overall performance are not sufficient to reflect the adaptivity of the persistence/flexibility tradeoff. This might be due to the nature of the tasks: The tasks that loaded on factor one (AUT, MAB, and 5PT) are characterized by the search for non-verbal concepts while the tasks that load on the second factor require more verbal resources, specifically, they require a search for words based on phonemics. This division might also explain why the RAT shows only weak factor loadings on both factors, while the task does require a search for words, this search is not based on phonemics but semantics (e.g., Mednick, 1962; Smith, Huber, & Vul, 2013).

3.2.1. Interim discussion

Our first two aims were to test whether clustering and switching generalized over tasks, which would be suggested by the trait-bias account, and whether total performance generalized over tasks, which could be supportive of the adaptivity account. From our findings we can conclude that clustering and switching do not reflect a metacontrol trait bias. Instead, the results speak more in favor of the adaptivity account, as total performance is at least partly task-independent. However, the lack of a relationship between some tasks and the two-factor exploratory factor analysis solution might indicate that a valid measure of adaptivity requires the consideration of either a broader range of tasks or a more specific range of tasks. As adaptivity should reflect individual variability that is independent of resources and situations, a heterogeneous sample of variables might be necessary to find this shared variance. However, it is thus important to balance the types of tasks very well, such that tasks cannot group together based only on shared variance that is unrelated to adaptivity per se. On the other hand, a more specific set of similar tasks might help measure adaptivity. In our results, the WPT, AUT, and 5PT are all fairly unrestricted tasks when it comes to instruction and task structure, in which participants fluently produce responses at their own tempo. The similarity in low task restrictions might allow participants to adapt the tradeoff between

Table 1
Spearman's Rho correlations (95% CI) between cluster size scores from all tasks (Holm-corrected for multiple comparisons).

	AUT		5PT		VST		MAB	
WPT	.092 (−.082, .260)	(n = 129)	−.006 (−.180, .168)	(n = 127)	.034 (−.214, .278)	(n = 64)	−.118 (−.352, .130)	(n = 65)
AUT			.042 (−.133, .215)	(n = 127)	.235 (−.012, .454)	(n = 64)	−.034 (−.275, .212)	(n = 65)
5PT					.174 (−.075, .402)	(n = 64)	.097 (−.155, .336)	(n = 63)
VST							.084 (−.168, .325)	(n = 63)

Note. WPT = word production task, AUT = Alternative Uses Task, 5PT = five point task, RAT = Remote Associates Test, VST = verbal search task, MAB = multi-armed bandit task. No correlations obtained statistical significance (p < .05).

Table 2
Spearman's Rho correlations (95% CI) between switching scores from all tasks (Holm-corrected for multiple comparisons).

	AUT		5PT		VST		MAB	
WPT	.140 (−.033, .306)	(n = 129)	.030 (−.145, .203)	(n = 127)	−.298 (−.507, −.057)	(n = 64)	−.184 (−.409, .063)	(n = 65)
AUT			.344*** (.180, .489)	(n = 127)	−.173 (−.402, .076)	(n = 64)	−.081 (−.319, .166)	(n = 65)
5PT					−.257 (−.473, −.011)	(n = 64)	.078 (−.173, .319)	(n = 63)
VST							−.008 (−.255, .241)	(n = 63)

Note. WPT = word production task, AUT = Alternative Uses Task, 5PT = five point task, RAT = Remote Associates Test, VST = verbal search task, MAB = multi-armed bandit task. *p < .05, **p < .01, ***p < .001.

persistence and flexibility according to their own resources to a similar extent. While the three tasks differ in the type of responses required (which is reflected in the EFA solution) they did share variance as expressed in significantly positive zero-order correlations and a positive correlation between the two factors in EFA. Moreover, of the three tasks that have more task restrictions, only the VST, which allows for at least some fluent production of responses according to the participant's verbal resources, correlates with fluent WPT.

3.3. Clustering and switching within tasks

Next, we tested whether clustering and switching together predicted total performance in all tasks (except the RAT, as clustering and switching measures were not available). In the WPT both clustering and switching were important predictors of total performance ($F(2, 126) = 160.6$, $p < .001$, $R^2 = .72$, $f^2 = 2.55$, Table 5). Both more clustering and more switching were related to better performance in terms of amount of words produced. Moreover, when corrected for the total score (as people who more rapidly produce words can both cluster and switch more), a negative partial correlation ($r_{partial} = -.657$, $p < .001$) between clustering and switching suggests that those who cluster more switch less and vice versa, a very intuitive association between the two. This indicates that participants indeed tradeoff clustering and switching to produce as many words as possible, biasing their search to be more persistent or flexible to produce more responses. Interestingly, switching was more strongly related to performance than clustering ($t(126) = 2.38$, $p = .019$), suggesting that this task benefits most from flexible cognitive control.

In the 5PT clustering and switching both positively predicted total performance as well ($F(2, 124) = 59.6$, $p < .001$, $R^2 = .49$, $f^2 = 0.96$, Table 5), again indicating that both clustering and switching are important for the fluent production of responses. As in the WPT, clustering and switching in the 5PT were inversely related when corrected for the total score ($r_{partial} = -.713$, $p < .001$), suggesting that clustering and switching could be considered two ends of a tradeoff. Unlike in the WPT, however, in the 5PT clustering and switching were equally important to produce as many designs as possible, as their correlation coefficients with performance did not differ significantly ($t(124) = 0.86$, $p = .391$).

In the AUT we again found that both clustering and switching positively predicted the total score ($F(2, 126) = 2493.0$, $p < .001$, $R^2 = .98$, $f^2 = 39.65$, Table 5), while being inversely related to each other when corrected for the total score ($r_{partial} = -.637$, $p < .001$).

Table 4
Factor loadings and communalities for oblimin-rotated exploratory factor analysis of total performance scores for all tasks.

	Factor 1	Factor 2	h^2
AUT	**.55**	.10	.35
MAB	**.52**	−.23	.23
5PT	**.51**	.19	.38
RAT	.33	.19	.20
VST	−.07	**.69**	.45
WPT	.23	**.59**	.51

Note. Factor loadings > .40 are in boldface. h^2 = communality. AUT = Alternative Uses Task, MAB = multi-armed bandit task, 5PT = five point task, RAT = Remote Associates Test, VST = verbal search task, WPT = word production task.

Table 5
Regression coefficients predicting total performance scores from clustering and switching per task.

	Standardized coefficients					
	β	*SE*	*t*	Low $CI_β$	High $CI_β$	*p*
WPT						
Cluster size	0.67	0.058	11.52	0.55	0.78	< .001
Switching	0.87	0.056	15.65	0.76	0.98	< .001
5PT						
Cluster size	0.67	0.075	8.94	0.52	0.82	< .001
Switching	0.75	0.075	9.99	0.60	0.90	< .001
AUT						
Cluster size	0.17	0.018	9.37	0.14	0.21	< .001
Switching	0.96	0.015	65.83	0.93	0.99	< .001
VST						
Cluster size	−0.18	0.146	−1.23	−0.47	0.11	.224
Switching	−0.64	0.146	−4.38	−0.93	−0.35	< .001

Note. SE = standard error, CI = 95% confidence interval bound, WPT = word production task, 5PT = five point task, AUT = Alternative Uses Task, and VST = verbal search task.

Similar to the WPT, total performance in the AUT seemed to benefit more from flexible than persistent processing, as the correlation coefficient between switching and performance was larger than that between clustering and performance ($t(123) = 15.72$, $p < .001$).

In the verbal search task only the regression coefficient of switching

Table 3
Spearman's Rho correlations (95% CI) total performance scores from all tasks (Holm-corrected for multiple comparisons).

	AUT		5PT		VST		MAB		RAT	
WPT	.301** (.135, .451)	(n = 129)	.412*** (.256, .547)	(n = 127)	.434** (.211, .614)	(n = 64)	.032 (−.213, .274)	(n = 65)	.212 (.039, .372)	(n = 127)
AUT			.427*** (.273, .560)	(n = 127)	.108 (−.142, .344)	(n = 64)	.094 (−.154, .330)	(n = 65)	.181 (0.007, .344)	(n = 127)
5PT					.220 (−.027, .442)	(n = 64)	.204 (−.046, .430)	(n = 63)	.151 (−.024, .317)	(n = 127)
VST							−.063 (−.306, .187)	(n = 63)	.122 (−.128, .357)	(n = 64)
MAB									.167 (−.084, .398)	(n = 63)

Note. WPT = word production task, AUT = Alternative Uses Task, 5PT = five point task, RAT = Remote Associates Test, VST = verbal search task, MAB = multi-armed bandit task. *p < .05, **p < .01, ***p < .001.

was significant and negative, indicating that those who switch more need less time to find 30 words ($F(2, 61) = 12.3$, $p < .001$, $R^2 = .29$, $f^2 = 0.40$, Table 5), in turn suggesting that only switching is beneficial for total performance on this task. This finding might deviate from the first three tasks because of the task constraints in the search for words: As participants have to find words consisting only of the presented letters it might be beneficial to switch well before depleting the letter set. Out of 14 possible letter sets, on average the participants used only 7.78 (mode = 6, median = 7) letter sets, indicating that switching sets was the most efficient manner to find 30 words as fast as possible. Similar to the WPT, 5PT and AUT clustering and switching were inversely related when corrected for total performance ($r_{partial} = -.670$, $p < .001$), moreover, switching was more strongly related to performance than clustering ($t(61) = 3.70$, $p < .001$).

The assumption of normally distributed errors was violated in the MAB and follow-up analyses showed very unstable results such that we could not reliably interpret results in the MAB.

3.3.1. Interim discussion

The results regarding the third aim, the question whether and how clustering and switching predicted total performance in each task, show that in similar, unrestricted tasks that require the fluent production of responses through cognitive search, both clustering and switching predicted higher total performance, while they were inversely related. These negative correlations suggest that clustering and switching indeed represent two ends of one dimension to be balanced on, but individuals might still prefer either clustering or switching for fluent production. However, the type of task seems to influence, and in fact constrain, the extent to which participants should choose to adapt their bias to persistent or flexible processing as shown by the different patterns of prediction. For example, as might be the case in the VST, task constraints might require all participants to adapt their metacontrol bias to the task constraints and switch more, disregarding their resources. This again offers support for the adaptivity account. Moreover, the contributions of clustering and switching to total performance show differences in the suitability of tasks to measure interindividual differences in cognitive control. For example, as switching was a better predictor of performance than clustering, the AUT seems well-suited to study interindividual differences in flexible cognitive control. On the other hand, to study interindividual differences in the balance between persistent and flexible control the 5PT might be more suitable as clustering and switching were equally important predictors of total performance.

3.4. General discussion

Our results suggest that the clustering and switching task scores that we have studied here are not reflective of individual metacontrol trait biases. Instead, the results suggest that these cognitive search tasks might be used as a starting point to study the adaptivity of a metacontrol state bias: to reflect to what extent individuals adapt their metacontrol bias to perform as optimal as possible considering the task restrictions and their task-related resources.

To prevent task restrictions and task-related resources from masking domain- and task-independent metacontrol biases and adaptivity, it is necessary to use either a well-balanced heterogeneous set of tasks or a more homogeneous set of tasks as well as to include measures of task-related resources (such as vocabulary). A heterogeneous set of tasks should be well-balanced in terms of task restrictions and requirement of resources, as adaptivity is expected to be domain- and task-independent. If tasks are not balanced in terms of either task restrictions or necessary resources, tasks might group together based on variance that is dependent on domain or tasks. Our results, however, suggest that tasks that are similar in terms of task restrictions—specifically, tasks that are fairly unrestricted such as the fluent production tasks in this study— allow participants to respond from their preferred metacontrol state bias. A more homogenous set of tasks to measure adaptivity as reflected by total performance should still consist of tasks that require a heterogeneous set of resources to prevent tasks from grouping together based on resource-related variance. Another possibility is that different behavioral measures might better reflect persistence or flexibility in cognitive control. For example, future research could study stopping rules in addition to or instead of clustering and switching (e.g., Harbison, Davelaar, & Dougherty, 2007). How participants terminate (exploitative) sampling from clusters might offer insight into their persistence in search (without any time restriction) and their reasons to stop sampling and switch to other clusters.

Importantly, however, metacontrol should be studied within a wider framework which, in addition to well-chosen behavioral tasks and measures aimed at studying the persistence/flexibility tradeoff, includes measures of neural mechanisms that might be related to metacontrol. While we aimed to do so in the current study, by including the dopamine proxies, the results did not allow us to draw substantive conclusions on the possible role of dopamine in metacontrol adaptivity (see also Supplementary material). Moreover, besides studying the role of dopamine or other neuromodulators (Avery & Krichmar, 2017; Doya, 2008; Goschke & Bolte, 2014) using more direct measures, recent insights have suggested other neural mechanisms of interest in cognitive control. For example, inter- and intraindividual differences in theta frequency network architectures might be related to the adaptivity of metacontrol (Zink et al., 2018).

Moreover, metacontrol and its adaptivity might emerge from or be related to interactions between multiple other cognitive processes or systems, such as working memory capacity and processing speed (Miyake & Shah, 1999; Unsworth et al., 2011). Placing the study of metacontrol in a broader framework of cognitive processes might offer insight into how intra- and interindividual differences in adaptivity come about. For example, Unsworth et al. (2011) showed how component processes of working memory performance and processing speed as well as vocabulary account for interindividual differences in performance on verbal fluent production tasks like the WPT. Related, Miyake & Shah (1999) summarize that many theories of working memory suggest that cognitive control emerges from interactions between knowledge and working memory or between subprocesses of the cognitive system. In their discussion, they recognize the importance of acknowledging domain-specific effects when studying cognitive control.

Summarizing, we suggest that metacontrol adaptivity might be an important process which should be taken into account when studying suggested domain-independent biases in tradeoffs in cognitive control (Hills, Todd, & Jones, 2015; Hills, Todd, Lazer, Redish, & Couzin, 2015). Individual differences in metacontrol adaptivity might explain the absence of zero-order correlations between clustering and switching tendencies in different tasks from different cognitive domains (as in the current study), the absence of a single latent factor to describe exploration/exploitation tendencies in recent research (Von Helversen, Mata, Samanez-Larkin, & Wilke, 2018), as well as the absence of zero-order correlations between different tasks that tap into executive functioning (as discussed in Miyake & Shah, 1999). To find a suitable measure of metacontrol adaptivity research should take place in broad framework that considers task restrictions, individual task-related resources, neurocognitive influences and the interaction between different cognitive processes.

Related to the exploratory nature of our study, we were faced with numerous limitations, such as an unbalanced set of tasks and a small sample size for some of the tasks. Future similar studies should consider using an improved balance in the set of tasks in order to study the domain- and task-restrictions independently from the variability that we assume to reflect adaptivity. Moreover, a larger sample size and a larger number of tasks would be particularly beneficial for inferring latent factors in exploratory factor analysis. Indeed, the fit measures that we found indicate that the model might have suffered from small

sample size. However, recent research shows similar conclusions in a well-powered study (Von Helversen et al., 2018). Although some of our non-significant results in Tables 1–3 in particular could have suffered from small sample sizes and type-II error, our conclusions would not change if these results would have been significant. Therefore, even in light of some small sample sizes, we consider our observations encouraging for further investigating and understanding individual differences in dealing with task demands adaptively.

Finally, it is important to note that the unrestricted fluent production tasks we considered in the present study were particularly suited to reveal adaptivity of metacontrol and these tasks may thus not have been particularly sensitive to reveal stable metacontrol biases. Accordingly, we do not take the present findings as evidence suggesting that such metacontrol preferences do not exist but, rather, we suggest that these findings indicate that they can be overcome to match situational demands.

4. Conclusions

We aimed to study cognitive metacontrol according to two approaches. We expected that our results would offer support for a metacontrol trait-bias account and/or a metacontrol adaptivity account. From the trait-bias account we expected clustering and switching tendencies to generalize over different types of cognitive search tasks. However, our data offers more support for the adaptivity account, as only total performance generalized (partly) over tasks, suggesting that participants adapt their balance between clustering and switching (or the metacontrol bias towards persistence or flexibility) from task to task, possibly according to task demands and individual resources. Finding good indicators of persistence and flexibility then requires the use of carefully chosen sets of tasks within a research framework that includes neurocognitive measures as well as measures of related cognitive processes from which metacontrol and metacontrol adaptivity might emerge or with which it might interact.

Importantly, we suggest that the degree to which people can engage in adaptive metacontrol adjustments varies substantially, and these individual differences seem rather consistent across a wide range of tasks in terms of overall performance. To summarize, our findings provide a starting point from where to study interindividual differences in metacontrol bias adaptivity.

Acknowledgements

We would like to thank our BSc students for collecting our data and all participants for taking part in our study. We would also like to thank Christian Beste and Thomas Hills for their valuable comments on a previous version of this manuscript. In addition, we thank Marieke Jepma for supplying us with her version of the multi-armed bandit task. Finally, we thank the European Research Council (ERC Advanced, # 694722) for funding our research.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.cognition.2018.10.001.

Aggregated data for this article can be found online at https://hdl.handle.net/10411/G369QS.

References

Akbari Chermahini, S., Hickendorff, M., & Hommel, B. (2012). Development and validity of a Dutch version of the Remote Associates Task: An item-response theory approach. *Thinking Skills and Creativity, 7*(3), 177–186. https://doi.org/10.1016/j.tsc.2012.02.003.

Ardila, A., Rosselli, M., & Bateman, J. R. (1994). Factorial structure of cognitive activity using a neuropsychological test battery. *Behavioural Neurology, 7*(2), 49–58. https://doi.org/10.3233/BEN-1994-7202.

Avery, M. C., & Krichmar, J. L. (2017). Neuromodulatory systems and their interactions: A review of models, theories, and experiments. *Frontiers in Neural Circuits, 11*(108), 1–18. https://doi.org/10.3389/fncir.2017.00108.

Bocanegra, B. R., & Hommel, B. (2014). When cognitive control is not adaptive. *Psychological Science, 25*(6), 1249–1255. https://doi.org/10.1177/0956797614528522.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items - applications to assessment of adaptive-behavior. *American Journal of Mental Deficiency, 86*(2), 127–137.

Colzato, L. S., Ozturk, A., & Hommel, B. (2012). Meditate to create: The impact of focused-attention and open-monitoring training on convergent and divergent thinking. *Frontiers in Psychology, 3*(116), 1–5. https://doi.org/10.3389/fpsyg.2012.00116.

Colzato, L. S., Sellaro, R., Samara, I., & Hommel, B. (2015). Meditation-induced cognitive-control states regulate response-conflict adaptation: Evidence from trial-to-trial adjustments in the Simon task. *Consciousness and Cognition, 35*, 110–114. https://doi.org/10.1016/j.concog.2015.04.012.

Colzato, L. S., Szapora, A., Lippelt, D., & Hommel, B. (2017). Prior meditation practice modulates performance and strategy use in convergent- and divergent-thinking problems. *Mindfulness, 7*(1), 152–159. https://doi.org/10.1007/s12671-014-0352-9.

Colzato, L. S., van der Wel, P., Sellaro, R., & Hommel, B. (2016). A single bout of meditation biases cognitive control but not attentional focusing: Evidence from the global-local task. *Consciousness and Cognition, 39*, 1–7. https://doi.org/10.1016/j.concog.2015.11.003.

Cools, R. (2015). The cost of dopamine for dynamic cognitive control. *Current Opinion in Behavioral Sciences, 4*, 152–159. https://doi.org/10.1016/j.cobeha.2015.05.007.

Cools, R., & D'Esposito, M. (2011). Inverted-U shaped dopamine actions on human working memory and cognitive control. *Biological Psychiatry, 69*(12), e113–e125. https://doi.org/10.1016/j.biopsych.2011.03.028.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441*(7095), 876–879. https://doi.org/10.1038/nature04766.

Doya, K. (2008). Modulators of decision making. *Nature Neuroscience, 11*(4), 410–416. https://doi.org/10.1038/nn2077.

Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: Reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning Memory and Cognition, 30*(2), 343–353. https://doi.org/10.1037/0278-7393.30.2.343.

Durstewitz, D., & Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to Catechol-O-Methyltransferase genotypes and schizophrenia. *Biological Psychiatry, 64*(9), 739–749. https://doi.org/10.1016/j.biopsych.2008.05.015.

Fischer, R., & Hommel, B. (2012). Deep thinking increases task-set shielding and reduces shifting flexibility in dual-task performance. *Cognition, 123*(2), 303–307. https://doi.org/10.1016/j.cognition.2011.11.015.

Gardner, E. G. S. (2008). *Strategy use on the Ruff Figural Fluency Test: An investigation of neuropsychological correlates and strategy exposure.* Idaho: State University.

Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98*(Pt 4), 611–625. https://doi.org/10.1348/096317907X173421.

Goschke, T. (2003). Voluntary action and cognitive control from a cognitive neuroscience perspective. In S. Maasen, W. Prinz, & G. Roth (Eds.). *Voluntary action: Brains, minds, and sociality* (pp. 49–85). New York: Oxford University Press.

Goschke, T., & Bolte, A. (2014). Emotional modulation of control dilemmas: The role of positive affect, reward, and dopamine in cognitive stability and flexibility. *Neuropsychologia, 62*, 403–423. https://doi.org/10.1016/j.neuropsychologia.2014.07.015.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Harbison, J. I., Davelaar, E. J., & Dougherty, M. R. (2007). *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (pp. 565–570). .

Hefer, C., & Dreisbach, G. (2016). The motivational modulation of proactive control in a modified version of the AX-Continuous Performance Task: Evidence from cue-based and prime-based preparation. *Motivation Science, 2*(2), 116–134. https://doi.org/10.1037/mot0000034.

Hefer, C., & Dreisbach, G. (2017). How performance-contingent reward prospect modulates cognitive control: Increased cue maintenance at the cost of decreased flexibility. *Journal of Experimental Psychology: Learning Memory and Cognition, 43*(10), 1643–1658. https://doi.org/10.1037/xlm0000397.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2008). Search in external and internal spaces. *Psychological Science, 19*(8), 802–808. https://doi.org/10.1111/j.1467-9280.2008.02160.x.

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: Priming exploration and exploitation across domains. *Journal of Experimental Psychology: General, 139*(4), 590–609. https://doi.org/10.1037/a0020666.

Hills, T. T., Todd, P. M., & Jones, M. N. (2015). Foraging in semantic fields: How we search through memory. *Topics in Cognitive Science, 7*(3), 513–534. https://doi.org/10.1111/tops.12151.

Hills, T. T., Todd, P. M., Lazer, D., Redish, D. A., & Couzin, I. D. Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences, 19*(1), 46–54. https://doi.org/10.1016/j.tics.2014.10.004.

Hommel, B., & Colzato, L. S. (2017). The social transmission of metacontrol policies: Mechanisms underlying the interpersonal transfer of persistence and flexibility. *Neuroscience & Biobehavioral Reviews, 81*, 43–58. https://doi.org/10.1016/j.neubiorev.2017.01.009.

Hommel, B., Colzato, L. S., Scorolli, C., Borghi, A. M., & van den Wildenberg, W. P. M.

(2011). Religion and action control: Faith-specific modulation of the Simon effect but not Stop-Signal performance. *Cognition, 120*(2), 177–185. https://doi.org/10.1016/j.cognition.2011.04.003.

Jepma, M., Beek, E. T. T., Wagenmakers, E. J., van Gerven, J. M. A., & Nieuwenhuis, S. (2010). The role of the noradrenergic system in the exploration-exploitation tradeoff: A psychopharmacological study. *Frontiers in Human Neuroscience, 4*(170), 1–13. https://doi.org/10.3389/fnhum.2010.00170.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation tradeoff: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience, 23*(7), 1587–1596. https://doi.org/10.1162/jocn.2010.21548.

Lu, J. G., Akinola, M., & Mason, M. F. (2017). "Switching On" creativity: Task switching can increase creativity by reducing cognitive fixation. *Organizational Behavior and Human Decision Processes, 114*(1), 29–41. https://doi.org/10.1016/j.obhdp.2017.01.005.

Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69*(3), 220–232. https://doi.org/10.1037/h0048850.

Miyake, A., & Shah, P. (1999). Toward Unified theories of working memory. In A. Miyake, & P. Shah (Eds.). *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 442–481). New York, NY: Cambridge University Press.

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills, 55*(3), 839–844. https://doi.org/10.2466/pms.1982.55.3.839.

Schmidt, C. S. M., Schumacher, L. V., Römer, P., Leonhart, R., Beume, L., Martin, M., ... Kaller, C. P. (2017). Are semantic and phonological fluency based on the same or distinct sets of cognitive processes? Insights from factor analyses in healthy adults and stroke patients. *Neuropsychologia, 99*, 148–155. https://doi.org/10.1016/j.neuropsychologia.2017.02.019.

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the Remote Associates Test. *Topics in Cognitive Science, 128*, 64–75. https://doi.org/10.1016/j.cognition.2013.03.001.

Stock, A. K., Steenbergen, L., Colzato, L., & Beste, C. (2016). The system neurophysiological basis of non-adaptive cognitive control: Inhibition of implicit learning mediated by right prefrontal regions. *Human Brain Mapping, 37*(12), 4511–4522. https://doi.org/10.1002/hbm.23325.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138–146. https://doi.org/10.1037//0894-4105.11.1.138.

Tucha, L., Aschenbrenner, S., Koerts, J., & Lange, K. W. (2012). The Five point Test: Reliability, validity and normative data for children and adults. *PLoS One, 7*(9), 1–11. https://doi.org/10.1371/journal.pone.0046080.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2011). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *Quarterly Journal of Experimental Psychology, 64*(3), 447–466. https://doi.org/10.1080/17470218.2010.505292.

Vannorsdall, T. D., Maroof, D. A., Gordon, B., & Schretlen, D. J. (2012). Ideational fluency as a domain of human cognition. *Neuropsychology, 26*(3), 400–405. https://doi.org/10.1037/a0027989.

Vik, P., & Ruff, R. R. (1988). Children's figural fluency performance: Development of strategy use. *Developmental Neuropsychology, 4*(1), 63–74.

Von Helversen, B., Mata, R., Samanez-Larkin, G. R., & Wilke, A. (2018). Foraging, exploration, or search? On the (lack of) convergent validity between three behavioral paradigms. *Evolutionary Behavioral Sciences, 12*(3), 152–162. https://doi.org/10.1037/ebs0000121.

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology. Adult, 23*(1), 29–34. https://doi.org/10.1080/23279095.2015.1004574.

Zink, N., Stock, A. K., Colzato, L., & Beste, C. (2018). Evidence for a neural dual-process account for adverse effects of cognitive control. *Brain Structure and Function*. https://doi.org/10.1007/s00429-018-1694-1 Advanced online publication.