Published in final edited form as:

Dev Rev. 2015 December 1; 38: 241–268. doi:10.1016/j.dr.2015.09.001.

A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children

Sabine Doebel[a,*] and **Philip David Zelazo**[b]

[a]Department of Psychology and Neuroscience, University of Colorado – Boulder, Boulder, CO, USA

[b]Institute of Child Development, University of Minnesota, Twin Cities, Minneapolis, MN, USA

Abstract

The Dimensional Change Card Sort (DCCS) is a widely used measure of executive function in children. In the standard version, children are shown cards depicting objects that vary on two dimensions (e.g., colored shapes such as red rabbits and blue boats), and are told to sort them first by one set of rules (e.g., shape) and then by another (e.g., color). Most 3-year-olds persist in sorting by the pre-switch rules, whereas 5-year-olds switch flexibly. We conducted a meta-analysis of standard and experimental versions of the task ($N = 69$ reports, 426 conditions) to examine the influence of diverse task variations on performance. Age, how the test stimuli were labeled for the child, emphasis on conflict in the verbal introduction of the post-switch rules, and the number of pre-switch trials each independently predicted switching on the standard DCCS, whereas pre-switch feedback, practice, and task modality did not. Increasing the relative salience of the post-switch dimension was associated with higher rates of switching, and, conversely, decreasing post-switch salience was associated with lower rates of switching, and under both kinds of manipulation performance continued to be associated with age. Spatially separating the dimensional values was associated with higher rates of switching, and it was confirmed that the degree of spatial separation matters, with children benefiting most when the dimensional values are fully spatially segregated. Switch rates tended to be higher in versions on which children were prompted to label the stimuli compared to when the experimenter provided labels, and lower when reversal instructions were used in conjunction with the standard task stimuli. Theoretical and practical implications for the study and measurement of executive function in early childhood are discussed.

Keywords

Cognitive control; Cognitive flexibility; DCCS; Executive function; Meta-analysis; Set-shifting

---

[*]Corresponding author. Department of Psychology and Neuroscience, University of Colorado – Boulder, 345 UCB, Boulder, CO 80309-0345, USA. sabine.doebel@colorado.edu.

Introduction

Being able to flexibly control our thoughts, actions, and emotions, especially in the face of conflicting habits and desires, is a key developmental achievement. Children are notorious for struggling with control, showing difficulty with simple tasks like sitting still and listening, raising their hand instead of blurting out thoughts, and flexibly switching from one activity to another. This skill, termed executive function (EF), develops markedly in the early childhood years (Carlson, 2005; Diamond, 2013; Zelazo et al., 2013), and can be described as the top-down, neurocognitive processes involved in goal-directed behavior, including inhibition of habitual or dominant responses, shifting between tasks or mental sets, and monitoring and updating information in working memory (Miyake & Friedman, 2012; Miyake et al., 2000). EF supports the development of a range of social and academic skills in childhood (e.g., theory of mind, math, and reasoning; Carlson & Moses, 2001; Hughes, 1998; Mazzocco & Kover, 2007; Richland & Burchinal, 2012) and also predicts numerous outcomes associated with success in adulthood (e.g., academic achievement, health, and wealth; Blair & Razza, 2007; Moffitt et al., 2011). Not surprisingly, deficits in EF can have wide-ranging negative consequences for adaptive functioning in children and adults (e.g., Bechara, Damasio, Damasio, & Anderson, 1994; Biederman et al., 2004; Pennington & Ozonoff, 1996). Given its central importance to healthy human functioning and development, gaining insight into the nature of EF and developing sensitive tools to measure it are key research goals.

One task that has been widely used in research on the development of EF in early childhood is the Dimensional Change Card Sort (DCCS) task (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006; see Fig. 1). In the *standard* DCCS (see Zelazo, 2006 for a complete description of the most commonly used and recommended procedure), children are shown two target cards (e.g., a blue rabbit and a red boat) and then presented with a series of test cards (e.g., red rabbits and blue boats) that they are instructed to sort according to one of the dimensions (e.g., by color). On each trial, the experimenter repeats the relevant rules, labels the test card by the relevant dimension, and gives the card to the child to sort. After several trials (typically 5 or 6), children are instructed to sort by the other dimension (e.g., shape). Three-year-olds tend to perseverate and continue to sort the cards according to the initial sorting rules. By contrast, typically developing 5-year-olds tend to switch flexibly. This task is often described as a measure of cognitive flexibility or set-shifting, and, like most measures of set-shifting, performance on the DCCS likely involves several aspects of EF, including working memory to maintain the rules in mind, and inhibitory control to suppress attention to the previously relevant dimension.

Over the past two decades, the DCCS has been used in over 150 studies and has been modified in theoretically motivated ways to examine cognitive mechanisms that underlie switching and EF. Developmental theories of EF that address the DCCS have focused on reflection and rule complexity (Zelazo, Müller, Frye, & Marcovitch, 2003), attentional inertia and inhibition (Kirkham, Cruess, & Diamond, 2003), representational redescription (Kloo & Perner, 2005), and working memory (Morton & Munakata, 2002).

On the Cognitive Complexity and Control-Revised account (CCC-r; Zelazo et al., 2003), successful switching requires the formulation and use of a higher-order rule for switching between dimensions (e.g., *If it's the color game, then the red ones go here and the blue ones go there; but if it's the shape game, then the rabbits go here and the boats go there),* which is achieved through developmental increases in reflection. Zelazo et al. (2003) examined performance on more than a dozen versions of the task to test theoretical predictions. For example, they tested the hypothesis derived from the CCC-r theory that a hierarchical rule is needed whenever rules are nested under different setting conditions (i.e., contexts that specify when it is appropriate to use a given rule), and found that children experienced difficulty switching in a modified version in which they needed to switch from sorting a single test card by shape to sorting it by color (e.g., sorting a green car with a green flower and switching to sorting it with a yellow car), and that children switched more on a version in which the rules conflicted but were not nested under different setting conditions (e.g., sorting a red and blue test card with a blue target card and then switching to sorting it with a red test card; Zelazo et al., 2003; Expts. 4 and 5).

By contrast, the Attentional Inertia account posits that children perseverate not because of the failure to construct a higher-order rule, but because they cannot inhibit the tendency to think of the stimuli in a particular way and continue to attend to the pre-switch dimension (Kirkham et al., 2003). This account has been tested by modifying the DCCS procedure such that the sorted test cards were left face-up in the sorting tray, which was hypothesized to increase the demands placed on children's fledgling inhibitory control. Results indicated that children were less likely to switch when the test cards were left face up. In another experimental condition, children were prompted to label the test cards themselves, which was expected to help children refocus attention to the new sorting dimension in the absence of a developed inhibitory mechanism, and, consistent with this hypothesis, children switched more often in this condition vs. when the experimenter labeled the card for them.

A related developmental theory of EF is the representational redescription account, which suggests that children's difficulty on the DCCS lies in their inability to appreciate that a single stimulus can be described differently from different perspectives (Kloo & Perner, 2005; Perner & Lang, 2002). One way this account has been assessed is by separating the depicted dimensional values on the test and/or target cards (e.g., a blue filled circle next to an outline of a banana, as in Kloo & Perner, 2005; see also Diamond, Carlson, & Beck, 2005). In this version, children do not have to change how they think about the stimuli in order to sort correctly (the blue thing is always the blue thing and the banana always a banana), and, consistent with the redescription account, children were more likely to switch on this version than on the standard task (Diamond et al., 2005; Kloo & Perner, 2005).

Others have argued that the ability to overcome perseveration depends on the development of sufficiently strong active representations that correspond to current goals. On the active–latent account (Morton & Munakata, 2002; Munakata, 1998), the ability to actively maintain information in working memory develops, and switching on the DCCS occurs when the actively maintained rules outcompete latent representations that correspond to the practiced pre-switch rules. One way this theory has been tested is by having children sort novel dimensional values during the pre-switch phase (i.e., unfamiliar colors or shapes), and familiar values during the post-switch phase (Yerys & Munakata, 2006). This was expected to result in children forming relatively weak latent representations of the pre-switch rules that young children’s active representations could compete with. Children tended to switch more in this condition vs. the standard task in which the dimensional values in both phases were familiar.

The extensive literature that has accumulated on the DCCS presents a unique opportunity to synthesize, clarify, and extend the findings of primary studies by examining the influence of a range of variations to both the standard and experimental versions of the task. Gaining further insight into the variations that influence switching has the potential to be both theoretically and practically significant. A meta-analysis of the DCCS can confirm theoretically relevant findings and also bring clarity in cases where there has been inconsistency both by aggregating across primary studies to reduce sampling error and increase precision in estimates (Hunter & Schmidt, 2004), and by including unpublished studies in an effort to address the bias to publish only positive results (Borenstein, Hedges, Higgins, & Rothstein, 2009; Ferguson & Heene, 2012; Lipsey & Wilson, 2001). It can also extend the findings of primary DCCS studies by classifying them in meaningful ways and examining patterns of performance across classes, which may, in turn, yield new findings in need of interpretation.

A meta-analysis of the DCCS can also provide practical insights that can inform the development of reliable, developmentally sensitive measures of EF. In addition to experimental manipulations designed to test hypotheses derived from specific theories, researchers have introduced variations to the standard task to address more practical concerns, such as the desire to maintain children’s interest in the task and ensure their understanding of the rules (e.g., how the experimenter labeled the test card, feedback on sorting during pre-switch phase, etc.). Determining the influence of such variations would provide new information on this widely used task, as well as new ideas about how changes to this and other EF tasks might be introduced to adjust EF demands. These insights can inform the development of measures to address various research aims, including measuring individual differences in EF, and also sensitively measuring change in EF across development and following interventions. In the following sections, we briefly describe key task variations that might be expected to influence performance on the DCCS.

Task variations in standard versions of the DCCS

Card labeling procedure

In the majority of studies involving the DCCS, the test card has been labeled by the relevant sorting dimension only (e.g., “Here is a *rabbit*. Can you put it where it goes [in the shape game]?”). In some studies, however, the test card has been labeled by both dimensions (e.g., “Here is a *red rabbit*”). This difference in the labeling procedure is subtle, but there are reasons to expect that it might affect performance. Prior research has found that providing children with labels that highlight task-relevant information facilitates their performance on cognitive tasks (Deák & Bauer, 1996; Müller, Zelazo, Hood, Leone, & Rohrer, 2004). Moreover, research with older children and adults indicates that using labels to highlight task-relevant information aids performance on more challenging task-switching paradigms

(Kray, Eber, & Karbach, 2008; Kray, Kipp, & Karbach, 2009). On the other hand, others have found evidence that relevant labels can *hurt* performance. One hypothesis, following from the active-latent account (Morton & Munakata, 2002; Munakata, 1998), is that labeling the pre-switch dimension strengthens active representations of the pre-switch rules, which in turn reinforces latent representations of those rules (Yerys & Munakata, 2006). Consistent with this suggestion, removing the relevant labels from the pre-switch phase has been found to improve switching compared to when relevant labels are provided in both phases (Yerys & Munakata, 2006).

Verbal emphasis on conflict between rules

Another subtle but potentially powerful task variation is the degree to which the contrast between the games is emphasized verbally at the introduction of the post-switch phase. In the standard DCCS protocol, the post-switch phase is introduced by announcing a new game, contrasting the two games explicitly, and then introducing and repeating the new rules (Zelazo, 2006; see Fig. 1 for the most commonly used version of the standard script). Other researchers have used language to further highlight the contrast between the conflicting rules in several additional ways, such as further emphasizing which game is *not* being played (e.g., "We're not playing the color game anymore, no way"; Munakata & Yerys, 2001), stating explicitly which rules are *not* to be used (Bohlmann & Fenson, 2005), and noting the novelty or distinctiveness of the new game (e.g., "The shape game is *different*"; Kloo & Perner, 2005; Munakata & Yerys, 2001). Highlighting the conflict between the games could lead to more flexible switching by providing additional opportunities for reflection on the hierarchical structure of the task and the incompatibility of the rules, which, according to the CCC-r theory, is critical to being able to effectively use rule representations to guide action. Perception of conflict has been theorized to trigger the activation of top-down control (Botvinick, Braver, Barch, Carter, & Cohen, 2001), initiating reflective reprocessing in prefrontal cortex (Cunningham & Zelazo, 2007), and individual differences in performance on the DCCS have been found to be related to the amplitude of a neural marker of conflict-detection (the N2 component of an event-related potential; Espinet, Anderson, & Zelazo, 2011).

Pre-switch feedback

In a number of studies using the standard DCCS, children have been provided with feedback (i.e., praise or verbal correction) on their performance during the pre-switch phase (e.g., Carlson & Moses, 2001; Kloo & Perner, 2005; Zelazo, Frye, & Rapus, 1996). Insofar as it provides explicit information about the correctness of sorting, pre-switch feedback might be expected to strengthen the formation of active representations of the pre-switch rules, which could in turn lead to the formation of stronger latent representations compared to conditions in which no such feedback is provided. The increased competitive strength of these representations against the active representations of the post-switch rules might then make switching more difficult. On the other hand, consistent with the CCC-r account, feedback might prompt more explicit awareness of the pre-switch rules (i.e., knowing *that* one is sorting by color rather than shape), which in turn could support subsequent flexible switching.

Number of pre-switch trials

Researchers have also varied the number of trials administered during which children sort according to the pre-switch rules. This has ranged from 1 to 15, with most researchers administering 5 or 6 trials. On the active–latent account, one would expect that each additional pre-switch trial would lead to incremental increases in perseveration, as the strength of latent rule representations is increased. This would be consistent with what has been found in the A-not-B task (Marcovitch & Zelazo, 1999) and with the notion of habit strength increasing with repeated behavior (Hull, Felsinger, Gladstone, & Yamaguchi, 1947).

Pre-test practice

Many studies using the DCCS have included a practice phase prior to having children complete the DCCS proper. This practice typically involved asking children to sort test cards that match target cards only on the relevant dimension and cannot be sorted based on the other dimension (e.g., practicing the color game by matching a blue truck to a blue bird and a red star to a red bird; Diamond et al., 2005). Although such practice procedures are typically brief, it is possible that they could influence children’s performance. If some children fail to switch on the DCCS due to a failure to understand the basic sorting rules, then one would expect practice to be associated with better performance. More interesting is the possibility that practice might strengthen representations of different sorting dimensions, like shape and color, which could in turn facilitate switching. On the CCC-r theory, practice could aid switching by providing children with sorting experience that could inform reflection on the task structure (i.e., that one can sort the same stimulus by color or by shape). Similarly, such practice could help children understand that the same stimulus can be thought of as a red thing or a round thing, consistent with the representational redescription account.

Task variations in experimental versions of the DCCS

Perceptual salience

Salience refers to those properties of a perceived stimulus that attract attention, and that “stand out” to a typical, healthy observer. For example, a bright red ball against a white background would be relatively salient to an ordinary perceiver, as would a sudden loud noise. Salient stimuli might be expected to exert an influence on attention and behavior in a relatively automatic, bottom-up fashion. In the standard DCCS, the pre- and post-switch stimuli and dimensions are presumed to be equally salient initially, so that children are not systematically biased toward one target card or toward sorting by one of the two dimensions. Numerous studies have manipulated properties of the pre- and/or post-switch stimuli in ways that might be expected to influence perceptual salience. For example, in one study flankers that depicted dimensional values corresponding to the post-switch rules were added to the test cards in the post-switch phase with the expectation that they would reduce the degree of conflict that children needed to resolve in order to overcome perseveration (Jordan & Morton, 2008; see Fig. 2a). Children switched more in this version compared to a control task with neutral flankers on the cards. In other studies, the salience of the pre-switch attributes in the post-switch phase has been reduced (e.g., by reducing the distinctiveness of the color values; Fisher, 2011a). In some cases, the relative salience of the two phases has been modified while testing other hypotheses. For example, Fisher (2011b) had children shift from sorting by shape (e.g., matching a mushroom with a similarly shaped open umbrella) to sorting by concept (matching an open umbrella with a closed umbrella), the latter being potentially weaker in perceptual salience. Examining salience modifications to the DCCS in the context of a meta-analysis can confirm and potentially extend the findings of primary studies by providing information on the degree to which such manipulations influence switching, and how they might affect the relation between switching and age.

Dimensional separation

Another way the DCCS has been modified is by spatially separating the dimensions on the test and/or target cards (see Fig. 2b for an example). Interest in dimensional separation and its influence on cognition has a long history (Garner, 1974; Garner & Felfoldy, 1970; Smith & Kemler, 1978; see Hanania & Smith, 2010, for a review) and recently has been explored in relation to EF (Diamond et al., 2005; Kloo & Perner, 2005; Zelazo et al., 2003). Separation of the dimensional values has generally been associated with better performance on the DCCS, and there are different interpretations about why this might be. For example, Kirkham et al. (2003) and Kloo and Perner (2005) have suggested that separating the values means that children do not have to change how they think about the stimulus, which they argue is the key difficulty accounting for young children's performance on this task. Another possibility is that separating the dimensional values reduces conflict by making it possible to attend to a single dimensional value with less interference from the competing dimensional value. Evidence for this kind of account has been found with adults on the Stroop task, where spatial separation of stimulus dimensions is associated with reduced conflict as evidenced by faster reaction times (MacLeod, 1991).

A meta-analysis can provide further insight about how separating the sorting dimensions affects children's performance. One way that separated versions of the task vary is in the amount of spatial distance between the dimensional values. In some modifications, the values are completely spatially segregated, such that the color value is presented on one half of the card and the shape value is presented on the other (e.g., Zelazo et al., 2003; Kloo & Perner, 2005). In other modifications, there is some degree of visuospatial contiguity (e.g., the color represented as a circle surrounding the shape, or the color forms the background on which the shape is imposed; e.g., Diamond et al., 2005; Kloo, Perner, Aichhorn, & Schmidhuber, 2010). On Kloo and Perner's representational redescription account, children might be expected to perform comparably across these two kinds of modifications since in both cases the need to consider a single stimulus from two perspectives is removed. On the other hand, if children experience more difficulty when the dimensions are not completely spatially segregated, this would indicate that some of their difficulty with the standard version could be attributed to factors other than representational inflexibility. For example, difficulty may vary with spatial proximity such that the more proximal the values, the more likely they will both receive attention, which in turn may activate both the relevant and irrelevant rules. The partially spatially separated versions would still be easier than the standard because children can focus attention more exclusively on the relevant dimension than in the standard version. A meta-analysis may be particularly informative on this question since this specific parameter was not manipulated within studies to test the influence of the degree of spatial separation, and thus the included conditions are more likely to be representative of the range of effect sizes.

Alternate versions

In addition to the experimental manipulations described above, many other modified versions of the DCCS have been created to test different theories (e.g., Zelazo et al., 2003; Kirkham et al., 2003; Perner & Lang, 2002). This affords the opportunity to further examine, at the meta-analytic level, whether performance on these versions differs from performance on the standard version. Specifically, a meta-analysis can clarify previous findings by reducing sampling error, increasing statistical power to detect an effect given a larger aggregate sample, and controlling for other factors that may have varied across studies, such as age. The supplemental material lists all versions of the DCCS considered for inclusion in the meta-analysis, with brief descriptions of the modification(s) made to the task stimuli and/or procedure. This includes versions that, while not identical, can be grouped together conceptually. For example, ‘hot’ versions of the DCCS, in which the motivational and/or emotional context of the task is enhanced (Zelazo & Carlson, 2012), have been created in a variety of ways, such as by using stimuli that signify potential rewards (Beck, Schaefer, Pang, & Carlson, 2011) or emotions (Qu & Zelazo, 2007). Meta-analyzing children’s performance on these different versions can confirm whether findings from individual reports hold when aggregated and compared to a more precise estimate of children’s performance on the standard task.

The current study

The aim of the current study was to conduct a series of meta-analytic tests to systematically examine the influence of diverse task variations on children’s performance on both standard and experimental versions of the DCCS in order to (a) confirm and extend theoretically relevant findings and (b) provide new information that could be useful for researchers seeking to create developmentally sensitive measures of EF to address a range of research goals. To pursue this aim, we conducted a series of weighted multiple meta-regression analyses on subsets of the data. First, we examined the effects of a range of variables (described above) on switching performance on the standard DCCS. Next, we examined performance on experimentally modified versions that varied dimensional separation and relative perceptual salience across task phases. Finally, we explored performance on several alternate versions of the task using both multiple regression and descriptive analyses.

Method

The unit of analysis in the current study was conditions within experiments, and the dependent variable was the proportion of children passing the post-switch phase within a condition. By using a single dependent measure for all conditions included in the meta-analysis, we were able to avoid the interpretational limitations associated with transforming reported inferential statistics (Borenstein et al., 2009).

Data collection

Aggregation of relevant published and unpublished studies was pursued using several strategies. First, thorough literature searches were conducted using databases such as PsycInfo and Google Scholar and using search terms that were broad enough to err on the side of being overly inclusive, such as Dimensional Change Card Sort (and all variants and abbreviations) and *child* (and variants) paired with terms such as *executive function*, *EF*, *inhibitory control*, *set shifting*, and *task-switching*. All relevant research reports were reviewed for additional citations and efforts were made to locate any article that appeared to report data from the standard DCCS or experimental variants. We also located studies by reviewing citations of early empirical reports involving the DCCS (e.g., Zelazo et al., 1996) as well as protocols and reviews (e.g., Zelazo, 2006) that were likely to be cited by researchers using the task. Abstracts from recent conferences that were likely to include empirical research in the area of cognitive development (e.g., meetings of the Society for Research in Child Development, Cognitive Development Society, Jean Piaget Society) were reviewed and authors of papers reporting DCCS data were contacted to solicit further information. Special effort was made to locate unpublished data because of the potential for publication bias to affect the reliability of the analysis (Borenstein et al., 2009; Littell, Corcoran, & Pillai, 2008; Rosenthal, 1979). Efforts to find relevant reports concluded in October 2012; no reports published after that time were included in the meta-analysis.

Inclusion criteria

Conditions had to meet a minimal set of criteria to be included in the meta-analysis in order to permit meaningful comparison with other conditions (Borenstein et al., 2009; Lipsey & Wilson, 2001; Littell et al., 2008). First, studies had to report the dependent variable of interest, the proportion of children passing the post-switch phase. Second, given that age was a primary variable of interest, the results had to be reported by age split by year (e.g., proportion passing reported for 3-year-olds and 4-year-olds separately). In studies where this was not done, the authors were contacted and a request was made for the data split by year of age. In cases where this information was not available, the study condition was excluded from the meta-analysis. Third, because the questions guiding the meta-analysis focused on typically developing children, we excluded the few studies that only reported data on atypical samples, such as children with autism. Fourth, because the meta-analysis focused on standard and modified versions of a specific task, it was critical to evaluate studies carefully to ensure that they retained essential properties of the DCCS so that meaningful comparisons across conditions could be made. For example, because the dependent variable in the meta-analysis was proportion switching (i.e., passing the post-switch phase), to be included in the meta-analysis, tasks had to involve sorting and have distinct pre- and post-switch phases with a change in rules in the post-switch phase. Tasks also had to involve at least two sorting locations, physical or virtual (as in the case of computerized versions). Fifth, to avoid dependency among conditions, which would compromise the analysis by violating statistical assumptions and increasing the risk of type I error (Borenstein et al., 2009; Lipsey & Wilson, 2001; Littell et al., 2008), repeated administrations of the same task to a single group were excluded from the analysis. Sixth, because the standard version of the DCCS was designed for use with typically developing 3- to 6-year-olds and because the vast majority of studies involve age groups within this range, conditions involving children aged 7 years and older were excluded.

Ninety-nine reports were considered for inclusion in the meta-analysis, yielding 615 conditions, of which 189 conditions were excluded given the criteria for inclusion listed above. The final sample included 69 reports and 426 conditions. See Table 1 for a list of all the studies included in the meta-analysis.

Coding of condition attributes

Conditions that met inclusion criteria were coded on the key independent variables described in the introduction. Some of these variables were readily extractible from reports (e.g., the number of pre-switch trials, publication year) and others were coded based on explicit criteria. Details regarding the coding scheme that was used are provided in the supplemental materials. In addition, three control variables were coded that had the potential to be confounded with the independent variables of interest: the *year* in which the study data were reported; the *criterion* used in a given study to judge whether the child had successfully switched/passed the task; and the *modality* in which the task was administered. Year of report was included to address the possibility that the initial effects found using the DCCS paradigm diminished over time, consistent with the decline effect (Schooler, 2011). Pass criterion was included to account for the possibility that children are more likely to pass the task when the criterion for passing is relaxed (e.g., 80% of cards sorted correctly vs. 88%). Task modality was included to account for possible variance due specifically to the format of the task (e.g., computer or tabletop). While there was no specific prediction regarding the computerized format, it was important to control for this variable because some experimental manipulations may have incidentally covaried with task modality. All conditions were coded by the first author and 20% were randomly selected and recoded by a research assistant with limited knowledge of the aims of the meta-analysis, the DCCS literature, or the hypothesized outcomes. Cohen's κ for all variables except degree of post-switch salience, separation, and conflict emphasis was 1.0, indicating perfect inter-rater reliability. Cohen's κ for the other three variables was between .91 and 1.0, indicating excellent inter-rater reliability on these variables. Discrepancies between coders was resolved through discussion of relevant cases.

Several variables of potential interest could not be coded and analyzed in the meta-analysis due to lack of variability or detail. For example, of potential interest is whether the procedure for restating the rules influences performance; however, we found that the vast majority of studies used the same procedure of repeating the rules on every trial. Other potential variables of interest such as which dimension children sorted by first (shape vs. color) and demographic variables like socioeconomic status, gender, and ethnicity, could not be coded due to insufficient detail provided in the reports. For example, most researchers reported that their samples were from middle- to upper-middle-class backgrounds; however specific information about the samples was generally not provided for the studies as a whole or by study condition.

Analytic approach

We conducted inverse-variance-weighted multiple regression analyses on four subsets of the data: standard versions; standard and salience versions; standard and separated versions; and standard and alternate versions. Each set of analyses included the following covariates: age, year, pass criterion, and task modality. The regression analyses were followed up with bootstrap analyses to confirm key findings. In each analysis, the independent variables were entered simultaneously, given the absence of an a priori rationale for entry in stages and to minimize the risk of type I error. This approach was also selected to maximize the informativeness of the analysis: given that independent variables in meta-analyses tend to be confounded to some degree (Lipsey & Wilson, 2001), by entering the variables into the model all at once, one can more clearly see whether a given independent variable predicts variance in the dependent variable, controlling for effects of all other independent variables.

Results

To satisfy the assumption of linearity that underlies regression, raw proportions were transformed into logits using the natural log (Lipsey & Wilson, 2001). Transforming the data rendered it appropriate for linear regression by increasing its range from negative infinity to positive infinity, as opposed to the restricted range between 0 and 1 that characterizes proportional data. While logits are not as intuitive as raw scores, they can be converted to back into proportions and odds ratios for interpretation.

Random effects inverse-variance-weighted multiple regression was used, as it was expected that the distribution of effect sizes (logits) would be heterogeneous, given both identifiable and unknown sources of between-condition variance. Random effects models use inverse variance weights that include variance components corresponding to subject-level sampling error and random effects between study conditions, while also assuming that some variance between study conditions is systematic and can be accounted for (Lipsey & Wilson, 2001). A model with a random variance component is preferable to a fixed effects model, which assumes a homogeneous distribution of effect sizes and, despite having more power to detect the influence of moderators, increases the rate of type I error. Heterogeneity was assessed using Cochrane's Q tests, which were conducted on each subsample of the data to be analyzed. In all cases, the presence of heterogeneity was confirmed, Qs > 311.43, ps < .001.

Another common issue in meta-analysis is multicollinearity, which reduces the accuracy of the regression coefficients, increases the risk of type I error, and makes it difficult to discern whether a given predictor is driving an effect. This was assessed separately for each subset of the data by computing Variance Inflation Factors (VIF). For all predictors, VIFs were less than 2, indicating the relative absence of multicollinearity.

Analysis of standard conditions

We first examined the relation between age and performance on the standard DCCS, and as expected, there was a robust correlation between age and the proportion of children passing the task, $r(190) = .66$, $p < .001$. Fig. 3 illustrates the marked age-related improvement in switching. As shown, and subsequently confirmed by our full regression model, by around 48 months of age, 50% of children pass the standard DCCS.

Next, we conducted a multiple regression analysis including, in addition to age and the previously described control variables, several independent variables of interest: labeling procedure, conflict emphasis, pre-switch feedback, number of pre-switch trials, and pre-test practice. Table 2 summarizes the results of the multiple regression and Table 3 shows the covariate adjusted logits and mean proportions for all predictors. The full model predicted that at 3.5 years of age, 41% of children pass the standard version, and that by 4.5 years 69% pass. The model accounted for 58.51% of the variance in performance, about 14.55% more variance than age accounted for on its own. Several variables predicted switching performance over and above age. The test card labeling procedure predicted performance ($p < .001$), such that children switched at higher rates when only the relevant dimension was labeled vs. when both dimensions were labeled. Conflict emphasis also predicted switching ($p < .02$), such that greater conflict emphasis was associated with higher switch rates. Finally, the number of pre-switch trials was a significant predictor of performance ($p = .03$), such that more pre-switch trials predicted lower rates of switching. Pre-switch feedback and pre-test practice, on the other hand, were not significant predictors. The three control variables also did not significantly predict switching.

Odds ratios for all significant predictors, controlling for all other variables included in the model, were computed to provide estimates of effect size (see Table 2) and can be interpreted as the change in the odds of passing the DCCS given a unit change in the level of a predictor variable. For example, the odds of passing the task are 10.35 higher for a 5-year-old vs. a 3-year old. Similarly, the odds of passing the task is 2.13 times higher if a child sorts test cards that are labeled by the relevant dimension vs. test cards that are labeled by both dimensions. Fig. 4 shows the relation between age and switching performance by labeling procedure. Covariate-adjusted means presented in Table 3 indicate that children who completed the standard DCCS with relevant labels are predicted to switch, on average, at a rate of 54%, whereas children who received the version in which both dimensions are labeled are predicted to switch at a rate of 35%.

Analysis of standard and salience conditions

Next we analyzed the standard version and all conditions coded as salience manipulations. In addition to the control variables, age and post-switch salience were included as independent variables. As summarized in Table 2, the analysis found main effects of age and post-switch salience. The model accounted for 38% of the variance in performance, with 4% attributable to relative post-switch salience. Results of a bootstrap analysis were consistent with and confirmed these results. We conducted two separate follow-up multiple regression analyses in which the salience variable was recoded into two dichotomous variables: high vs. neutral, and neutral vs. low relative post-switch salience. Both of these variables were significant predictors of performance over and above age, ($\beta = .19$, $p < .001$, and $\beta = -.13$, $p < .01$, respectively). Fig. 5a illustrates the relation between age and performance by degree of post-switch salience.

Analysis of standard and separated conditions

This analysis focused on conditions that varied in the degree of spatial separation of the dimensional values on the test and/or target card, and included the three control variables, age and degree of separation as predictors. As summarized in Table 2, age and degree of separation significantly predicted switching, and a secondary analysis using bootstrapping methods confirmed these findings. To further explore the influence of degree of separation on switching, we conducted follow-up analyses in which degree of separation was recoded into three variables to facilitate pairwise comparisons between partial separation, full separation, and full integration (i.e., standard) conditions. These analyses found that only the difference between fully separated and integrated conditions was significant, $\beta = .18$, $p < .001$. There was a non-significant pattern indicating higher rates of switching in partially separated vs. fully integrated conditions, $\beta = .08$, $p = .14$. Fig. 5b illustrates the relation between age and switching by degree of separation.

Analysis of alternate versions

In our final set of multiple regression analyses, we examined whether particular alternate versions of the task (beyond those examined so far) would predict variance in switching. This was undertaken by comparing each of six alternate versions for which there were at least eight conditions coded to the standard version in individual multiple regression analyses. Table 2 summarizes these analyses (see the supplemental materials for brief descriptions of these versions) and Table 4 presents the covariate-adjusted means of these alternate versions. A preliminary analysis including all versions and the standard conditions found that age and modality were significant predictors of switching. The individual regression analyses therefore included age and modality in addition to the version under analysis. Three alternate versions of the DCCS predicted significant variance in switching: unidimensional reversal, target characters instead of cards, and goal neglect (in which some post-switch test cards completely matched one target). In addition, there were marginally significant findings involving the child label and bidimensional reversal versions. As indicated in Table 2, the version in which children are prompted to label the test card as opposed to the experimenter providing the label tended to be easier than the standard version. A follow-up regression analysis that included labeling procedure to control for the possibility that children performed better on the child label version because only the relevant dimension was cued (e.g., in the shape game asking what shape was depicted on the test card) found a similar pattern. The bidimensional reversal version, in which the stimuli include irrelevant dimensional values and reversal shift instructions are provided (e.g., switching from “same” to “silly” game), tended to be more difficult than the standard version.

Exploratory analyses suggest the characteristic age-performance relation might have been influenced in these versions that significantly differed from the standard task. Controlling for modality, age and proportion passing the task were significantly correlated for both the unidimensional reversal and target character versions, $r_{partial}(11) = .74$, $p < .01$ and $r_{partial}(10) = .66$, $p < .05$, respectively; whereas there were no significant correlations between age and performance on the observation version, $r_{partial}(14) = .21$, $p = .45$, and $r_{partial}(24) = .23$, $p = .26$. This information could not be computed for the goal neglect version since all conditions were administered on the computer, but the simple bivariate correlation showed no significant age-performance relation, $r(12) = .11$, $p = .73$.

Finally, we conducted descriptive analyses, reported in Table 5, on the remaining alternate versions for which there were too few conditions to submit to multiple regression analysis (see the supplemental materials for descriptions of each of these versions).

Discussion

The current study confirms and extends our knowledge about task variations affect switching performance on the DCCS, a widely used measure of EF in childhood. In addition to confirming that age is a robust predictor of switching on the standard version, labeling the test card by the relevant dimension and verbally emphasizing the conflict between the pre- and post-switch games were both positively associated with switching, whereas the number of pre-switch trials was negatively associated with switching. Verbal feedback provided during the pre-switch phase, pre-test practice, and task modality were not significant predictors of switching on the standard task. Children also tended to switch more when the post-switch dimension was more salient than the pre-switch dimension, and tended to switch less when the post-switch dimension was less salient than the pre-switch dimension, confirming the findings of primary studies. Similarly, we confirmed that spatial separation of the dimensional values is positively related to switching, and extended prior research by showing that this effect was stronger when values were fully spatially segregated on the test card. We also confirmed that several alternate versions of the DCCS were more or less challenging than the standard version, and found evidence suggesting children tend to switch more when they label the test card themselves vs. when an experimenter provides the label. Finally, we found new evidence that children tend to have more difficulty switching on the standard, bidimensional task when the standard instructions are replaced by reversal instructions. We discuss these findings in relation to developmental theories of EF and then consider their implications for the development of sensitive measures of EF in childhood.

Task variations influencing performance on the standard DCCS

That age was a robust predictor of performance in all analyses further confirms the rapid development of EF that occurs during the early childhood years. The magnitude of this effect indicates a striking increase in switching between three and five years of age. An important goal for future research is to further explore how much this transition represents key changes in cognitive flexibility, and EF more broadly, at this age, and how well these patterns generalize to more diverse samples.

That children switched more when the test card was labeled by the relevant dimension vs. when it was labeled by both dimensions is consistent with previous findings that goal-relevant labels aid cognitive flexibility. A plausible interpretation of this finding is that relevant labels help children shift their attention to the relevant dimension in the absence of an endogenous mechanism to support EF. This interpretation does not appear to challenge any of the extant developmental theories of EF; however, it might be argued that children do not perform better in the relevant label conditions because of the presence of the relevant labels, but rather they do so because of the *absence* of the label referring to the conflicting dimension. This would be consistent with findings that removing relevant labels from the pre-switch phase improves switching (Yerys & Munakata, 2006), as well as findings that referring to both dimensions when asking children where the test cards go in the post-switch phase results in more incorrect responses than when referring only to the relevant dimension and diminishes the dissociation between knowledge and action (Munakata & Yerys, 2001). On the active–latent account, labeling the conflicting dimension in the pre-switch phase might not be expected to have any influence on performance, but doing so in the post-switch phase would be expected to further activate latent representations of the pre-switch rules. There are, however, other findings suggesting relevant labels aid performance on other tasks and in other age groups (e.g., Deák & Bauer, 1996; Kray et al., 2008; Kray et al., 2009; Müller et al., 2004).

If it is indeed true that relevant labels help children switch on the DCCS, this would raise the possibility that providing such labels might also help to foster the development of EF, which would be consistent with the notion that EF has social origins and that external verbal support for coordinated behavior fosters the development of children's ability to use their own speech to achieve the same ends (Winsler, Fernyhough, & Montero, 2009; Vygotsky, 1935/1962). Research with older children and adults is consistent with the idea that self-directed speech is used to cue goal-relevant information and that this ability develops (Chevalier & Blaye, 2009; Emerson & Miyake, 2003; Kray et al., 2008; Miyake, Emerson, Padilla, & Ahn, 2004). Future research can examine whether labels that support switching in the moment may also lead to longer term gains in flexibility.

Verbal emphasis on the conflict between the pre- and post-switch phases also predicted higher rates of switching. This finding is consistent with the possibility that emphasizing the conflict between the two games interacts with children's fledgling capacity for reflection, helping them to represent the hierarchical rule structure of the task and detect conflict (Botvinick et al., 2001; Zelazo, 2004). Alternatively, such language, insofar as it emphasizes which game is *not* being played, and further highlights the new rules, may influence sorting behavior by helping orient attention away from the pre-switch dimension in the absence of an inhibitory control mechanism (Doebel & Zelazo, 2013; Kirkham et al., 2003). Another possibility is that the additional contrastive language provides children with more opportunity to recognize that a change is occurring and that something different is coming up, which may help children to clear the old rules from working memory and update to the new rules. The ability to update to a new task set is known to play a role in EF in adults (Miyake & Friedman, 2012; Miyake et al., 2000) and recent work has suggested a specific role for updating in performance on the DCCS (Chatham, Yerys, & Munakata, 2012). This finding thus raises interesting possibilities concerning how linguistic input may interact with emerging developmental mechanisms to support EF that can be tested in future research.

The number of pre-switch trials was also a significant predictor of performance, over and above labeling and conflict emphasis, and was negatively associated with switching on the standard version of the task. This is consistent with the effects of experience on habit strength (Hull et al., 1947) and findings from other childhood EF tasks, notably the A-not-B task (Marcovitch & Zelazo, 1999). In terms of the active–latent account, adding more pre-switch trials may strengthen latent representations of pre-switch rules, which in turn increases the demand on working memory (Morton & Munakata, 2002). Importantly, while our findings suggest that the number of pre-switch trials influences performance, we do not know if this is represents a linear effect given the limited variability in number of pre-switch trials in the extant studies, with the vast majority of studies administering 5 or 6 pre-switch trials. The CCC-r theory, for example, predicts that beyond a moderate number of pre-switch trials, additional trials could result in improved performance, as seen on the A-not-B task in both infants and toddlers (Marcovitch & Zelazo, 2009) and in adults in the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948; Grant & Cost, 1954), by providing opportunities for reflection on the task structure, including noticing the other relevant dimensions. Future research can examine the nature of this finding.

Some variations in procedure did not appear to matter to performance on the standard task. Neither verbal feedback during the pre-switch phase nor pre-test practice were significant predictors of switching. Some researchers may have used feedback during the pre-switch phase with the assumption that it would motivate children to continue the task or even further consolidate the prepotency of the pre-switch rules, and others may have avoided using it in light of concerns about how it might influence children’s sorting behavior and invalidate the task. Our findings indicate that providing feedback during the pre-switch phase may not have any influence on performance. Similarly, pre-test practice has been used in EF tasks to ensure that children understand what is being asked of them; however, our findings suggest that children’s performance does not vary depending on whether an opportunity to practice is provided or not.

Task modality (computer vs. table-top) also did not significantly predict switching on the standard version of the task. It was, however, a significant predictor in the analysis that included alternate versions, perhaps because many computerized experimental manipulations were modifications in which the task was easier than the standard version. Computerized versions of table-top tasks might be expected to be more difficult for children given that the stimuli are less concrete, but this analysis suggests that, for the DCCS at least, this format is not more difficult for children.

Task variations influencing performance on modified versions of the DCCS

The analysis of experimental conditions revealed that the relative salience of the post-switch dimension predicted switching on the DCCS such that switch rates tended to be higher in conditions in which the relative salience of the post-switch dimension was high compared to “neutral” conditions in which the salience of the two dimensions was balanced. The effect size was comparable to that found for labeling the relevant dimension vs. both, and for increasing age by one year. Conversely, switch rates were lower in conditions in which the relative salience of the post-switch dimension was low compared to neutral conditions. The effect size corresponding to this difference was comparable in magnitude to the effect size corresponding to the difference between relatively high post-switch salience vs. neutral conditions. These results confirm and extend the findings of primary studies using specific salience manipulations (e.g., Fisher, 2011a; Honomichl & Chen, 2011; Jordan & Morton, 2008), showing that they hold when considering conditions that were not devised as salience manipulations but which could be reliably classified as such.

Some have interpreted such findings as evidence against accounts like representational redescription and CCC-r because core aspects of the task are presumed to be held constant (e.g., rule complexity, clash between the target and test cards; Jordan & Morton, 2008) and thus children should still fail to switch if they lack the ability to reflect and formulate a higher-order rule or view a stimulus from different perspectives. Although such findings do indeed weigh against “all-or-nothing,” stage-like accounts, it is debatable whether such modifications leave all relevant aspects of the task unchanged. For example, children may perform better when the post-switch dimension is relatively salient because this increases the activation of the corresponding rules in a bottom-up fashion, and children may then use these rules to guide their behavior even in the absence of reflection (Zelazo et al., 2003). The CCC-r account seeks to explain perseveration in the face of conflict (i.e., in the absence of such bottom-up support that reduces such conflict), by positing that what develops is children’s ability to reflect, notice the task structure, and represent it by integrating the simple rules under a higher-order action rule (Zelazo et al., 2003). Moreover, according to the CCC-r theory, the ability to reflect is not something that develops in a stage-like fashion. Rather, what increases with age and experience is the probability of reflection in a given situation (e.g., Marcovitch & Zelazo, 2009), and at all ages, a range of factors (e.g., motivation, fatigue, degree of conflict) can influence that probability that one will reflect at opportune moments (Zelazo & Doebel, 2015; Zelazo et al., 2003). For example, if the salience of the pre-switch dimension is enhanced via the addition of flankers that cue that dimension, children may be vulnerable to being lulled into responding by that dimension vs. reflecting on what game is being played and fail to switch or make occasional errors.

In sum, while it is clear that salience manipulations do influence children’s behavior on the DCCS, it is unclear how these variations influence the EF demands inherent in the standard task. They may be attenuated or enhanced, or obviated completely. Or, insofar as the DCCS measures a complex skill that is subserved by more basic cognitive capacities (Miyake & Friedman, 2012; Miyake et al., 2000), it could be that these changes to the task reduce or remove demands on some of these basic capacities but not others. For example, one intriguing possibility is that increasing the salience of the post-switch dimension removes the need to formulate a higher-order rule or redescribes the stimulus to overcome the conflict, but retains the need to maintain the post-switch rule in working memory to a degree that it can compete with the prior activation of the pre-switch rule, consistent with the active–latent account. This could explain why one continues to see an age-performance relation under such modifications. These questions are worth investigating further, both to gain insight into underlying mechanisms and also to gain the clarity needed to address practical considerations related to the measurement of EF in children. We discuss the latter in detail when we return to these results to consider the implications of our findings for the creation of developmentally sensitive measures of EF.

Degree of dimensional separation was found to aid switching, and partial separation (e.g., shapes and colors are not integrated but are contiguous or overlapping) appeared to be less facilitative than full separation (e.g., shape on one half of the card and color on the other half). This finding is consistent with the interpretation that some of children’s difficulty on the standard version may be attributed to interference arising from the visual processing of features of the stimulus that activate the pre-switch rules. That is, when the features are partially separated, it is easier to ignore the irrelevant features vs. when they are integrated, and it is much easier still when they are completely segregated to different halves of the test card. Moreover, the age-performance relation remains when the dimensional values are partially separated, with older preschoolers tending to switch at higher rates, whereas on the fully separated version the age-performance relation is completely diminished. This suggests that when the stimuli are contiguous or overlapping, the features corresponding to the conflicting dimension are still processed (or have a higher probability of being processed) and thus EF is needed to resolve the conflict. On the other hand, it could be that some factor other than conflict and EF demands explains why children do better when the dimensional values are fully spatially separated vs. partially, and why performance in partial separation conditions is associated with age, while performance on the fully separated conditions is not. Future research can test these different possibilities.

Patterns of performance on alternate versions were generally consistent with the findings of primary studies; however two findings in particular call for further examination. We found that children tended to switch at higher rates when they were prompted to label the test card for themselves vs. when the test card was labeled by the experimenter, consistent with the primary findings of Kirkham et al. (2003) and contrasting with the findings of Müller et al. (2008). We highlight a couple of reasons why this finding may have emerged in the meta-analysis. First, meta-analyses have more power to detect effects because sample sizes are larger and sampling error is reduced. For example, our estimate of performance on the standard version is more precise than what would be obtained in a single study. Second, in meta-analyses it is possible to control for covariates such as age that can make comparing primary studies difficult.

That children switch at higher rates when they label the test card has been interpreted as support for the Attentional Inertia account, such that labeling helps children disengage from thinking about the stimuli in terms of the pre-switch dimension and reorient to the post-switch dimension in the absence of a well-developed inhibitory mechanism. Integrating this with the finding that relevant labels appear to aid switching, prompting children to label the card on their own might be a more potent intervention that engages attention more strongly than when the experimenter labels the card. However, this finding is also consistent with other developmental accounts that allow that attention could be redirected in a bottom-up fashion via labels or other salient aspects of the task. Another possibility is that the experimenter's prompt to the child to label the card may encourage reflection, which in turn may allow children to construe the higher-order structure of the task and resolve the conflict, consistent with the CCC-r theory (Müller et al., 2008; Zelazo, 2004).

The meta-analysis also indicates that children tended to perform worse on the bidimensional reversal version of the task, in which the task stimuli include variation on two dimensions, but only one dimension is relevant because children were simply instructed to reverse their sorting pattern (e.g. switching from the "same" to "silly" game) as opposed to hearing the standard instructions that refer to a dimensional shift. In some cases, the task stimuli were identical with that of the standard task (i.e., bidimensional with each target card conflicting with test cards on one of two dimensions; Brooks, Hanauer, Padowska, & Rosman, 2003a; Kloo, Perner, Kerschhuber, Dabernig, & Aichhorn, 2008). Children's relatively poor performance on the bidimensional reversal contrasts with their relatively good performance on the unidimensional reversal, which also uses the "same" to "silly" game instructions but with stimuli that only vary on one dimension. One possibility is that children notice the changing values on the second dimension and are confused by their presence because the experimenter does not refer to them, but instead refers to "same/silly." If confusion is a key factor influencing performance on this version, children might be expected to make spontaneous comments and queries about the second dimension. It is also possible that some children spontaneously conceptualize the pre-switch game in terms of the dimensions or dimensional values being sorted (e.g., shape game, or rabbits go with rabbits and boats go with boats), especially given that the test and target cards do not fully match and thus matching must be selective according to one dimension or dimensional value. If so, then "same/silly" formulation may confuse them. In the unidimensional reversal version, which we discuss in the following section, this problem goes away because the test and target cards fully match so the "same" game instructions are more sensible.

Confirming and extending the findings of primary studies (e.g., Kloo et al., 2008; Perner & Lang, 2002), the unidimensional reversal and target character versions were generally associated with higher rates of switching than other versions with the adjusted mean passing rates of 81% and 85%, respectively, and performance improved with age. This raises the possibility that these versions, while different from the standard task, may still tap age-related change in EF by removing some but not all of the conflict. On the other hand, the finding that these versions are easier than the standard version has been interpreted as supporting the representational redescription account (Perner & Lang, 2002). In the unidimensional version, children match test and target cards that vary on one dimension (e.g., sorting red suns with red suns and red cars with red cars) and then are instructed to switch to sorting the test card in the opposite way (e.g., sorting red suns with red cars, etc.). Similarly, the target character version, in which the target cards are replaced by characters such as Donald Duck and Mickey Mouse, requires children to sort test cards one way, in accordance with the characters' ostensible preferences, and then switch to sorting the cards in the opposite way in according with their new preferences. On the representational redescription account, these versions are easier because they provide additional cues to the child that the test card is to be viewed differently in each game. Another possibility, however, is that the unidimensional reversal version does not require shifting between two sets of conflicting rules but rather modifying the initial rule into its opposite (a possibility discussed in Perner & Lang, 2002). This interpretation is consistent with findings that when given feedback during the post-switch phase on the standard task some children essentially play an opposites game and do not truly switch between rule sets, as evidenced by the fact that they sort cards that fully match a target with the opposite target instead (Chatham et al., 2012). It is also consistent with neuroscientific evidence that the neural correlates of performance on reversal tasks is different from tasks known to tap aspects of EF like working memory, inhibition and set-shifting (Cools, Clark, Owen, & Robbins, 2002; Hampshire & Owen, 2006). An alternative interpretation of performance on the target character version is that the pre-switch dimension is less salient given that the test and target cards do not physically match on any dimension, and thus switching may occur in the absence of a robust EF mechanism.

Also consistent with published findings, children tended to switch at lower rates on the goal neglect version compared to the standard task (Marcovitch, Boseovski, & Knapp, 2007; Marcovitch, Boseovski, Knapp, & Kane, 2010). In this version, some of the post-switch test cards match the target on both dimensions, which reduces the need to actively maintain the goal on those trials. This is expected to lead children to lower their active maintenance of the rules across all trials and consequently make errors when sorting test cards in which there is conflict because they fall back on the prepotent pre-switch rules. No age-performance relation was found on this version, which could suggest that differences beyond age are more critical to performance on this task, such as individual differences in working memory (Marcovitch et al., 2007, 2010).

No difference was found between performance on the standard task and versions in which children watched another agent – a person or puppet – sort cards in the pre-switch phase, and then were asked to sort the cards according to a new rule in the post-switch phase, or were asked to judge the agent's post-switch sorting behavior. This finding is consistent with notion that EF failures as indexed by the DCCS are likely due to representational inflexibility rather than problems inhibiting a prepotent motor response (Jacques, Zelazo, Kirkham, & Semcesen, 1999). However these versions may have varied in nontrivial ways procedurally – for example, in some cases the agent and/or the experimenter provided feedback on the agent's sorting (e.g., Moriguchi, Lee, & Itakura, 2007). Nevertheless, followup regression analyses did not reveal any difference between switching performance on conditions with or without such feedback and performance on the standard version.

The results of descriptive analyses on the remaining alternate versions should be interpreted conservatively because covariates could not be controlled for statistically and the small number of conditions increases the likelihood that heterogeneity influenced the estimates. With these caveats, we note that hot versions of the DCCS did not appear to be more challenging than the standard task, in contrast to some theorizing about the developmental trajectories of hot vs. cool EF (Zelazo & Carlson, 2012). One possibility is that in some cases enhancing the motivational significance of the task actually facilitated performance, consistent with current reward-based models of cognitive control, wherein effort is a function of anticipated costs and benefits of engaging control (Botvinick & Braver, 2015). Further research is needed to clarify precisely how motivationally and emotionally significant situations influence young children's ability to exercise EF on the DCCS and other child EF tasks. Other findings appear generally consistent with what has been observed in primary studies. For example, performance on the negative priming version did not differ from performance on the standard task, which is consistent with the notion that part of children's difficulty on the standard task is the failure to engage attention to dimensional values that were irrelevant and ignored during the pre-switch phase (e.g., attending to red and blue after sorting by shape; Müller, Dick, Gela, Overton, & Zelazo, 2006; Zelazo et al., 2003).

To summarize the implications of the current meta-analysis for developmental theories of EF, the reported findings both clarify and extend our knowledge of how a range of task modifications influence switching performance on the DCCS and raise new questions for future research. We confirmed that children's performance on the standard DCCS improves dramatically between 3 and 5 years of age; future research can examine the degree to which this pattern may be moderated by demographic variables such as socioeconomic status. We confirmed that labeling procedures influence children's tendency to switch flexibly; research examining the conditions under which labels help or hinder flexibility can provide more insights into developmental mechanisms. Verbally highlighting the conflict between the games appeared to help children switch; future research can examine the nature of this novel finding. More pre-switch trials predicted lower switching rates, but whether this relation is linear or more complex is not known. The findings with respect to key alternate versions of the DCCS raise new questions about developmental mechanisms underlying changes in EF in childhood. Specifically, through what mechanisms do manipulations like spatially separating the dimensional stimuli, increasing post-switch dimensional salience, replacing the target cards with characters, prompting the child to label the test card, and replacing the standard instructions with reversal instructions, influence switching? And what, if any, implications do the age-performance correlations seen on some of these modified versions suggest about underlying mechanisms? The meta-analysis raises more questions than it answers regarding the best theoretical interpretation(s) of performance on the DCCS, and highlights the extent to which multiple theories can account for the same findings. We hope these findings will inspire new efforts to find novel ways to test and further develop theories that avoid some of the interpretational challenges emphasized here and advance our knowledge of developmental mechanisms of EF.

Implications for the measurement of EF in childhood

The meta-analysis also provides information of practical value that can inform efforts to create sensitive measures of childhood EF to address various research aims, including measurement of individual differences and developmental change in EF, and change in EF induced by interventions. A key general finding is that a range of modifications to the DCCS task procedure and stimuli make the task more or less challenging for children while in many cases seeming to preserve the age-performance relation. This interpretation is made with the caveat that it is possible that in some of these cases changes to the task introduce and place new demands on other cognitive skills that develop with age and that the relation does not necessarily reflect age-related improvements in EF. Table 6 summarizes the modifications to the DCCS that our findings suggest enhance, reduce, or do not seem to affect its difficulty, which we suggest may operate by decreasing, increasing, or not modifying the demands placed on EF, respectively. Such modifications may also function similarly in other childhood EF tasks. For example, the DCCS and other rule-based child EF tasks may be made more challenging by reducing the perceptual salience of features of the stimuli that are consistent with the relevant rules (e.g., using novel shapes or colors, more abstract dimensions or relational stimuli), and, conversely, increasing the perceptual salience of features of the stimuli that highlight the conflicting rules (e.g., familiar shapes or colors, concrete or familiar dimensions, flankers that cue the conflicting rules). EF demands can be increased by removing labels or symbols that cue or prime the relevant rules, removing contrastive language that highlights conflict between competing task rules, and increasing the number of trials on which children must respond to goal-conflicting aspects of a stimulus. Moreover, our results suggest the DCCS and other child EF tasks can be made *less* demanding on EF by using familiar labels to denote goal-relevant aspects of a stimulus and encouraging children to use those labels themselves, enhancing the perceptual salience of goal-relevant information, separating the dimensional values, and using contrastive language to emphasize conflict between competing rules. Modifying child EF tasks in these ways can address common issues in the measurement of cognitive change by attenuating floor and ceiling effects and extending the age range that can be measured by a single task. Administration of EF task batteries that include assessments of different components of EF (e.g., Weibe et al., 2011; Willoughby, Blair, Wirth, & Greenberg, 2012) can be challenging, as relatively few tasks are appropriate for use across age groups given how rapidly EF develops during this age.

Our findings also suggest that some task features can be modified parametrically to create graded measures that are sensitive to individual differences and allow researchers to assess childhood EF with more precision. For example, labeling procedures can be adjusted such that children are first prompted to label aspects of a stimulus that correspond to the current rules, followed by the experimenter providing these labels for the child, and finally the removal of labels altogether. Likewise, the findings involving the separated conditions suggest that separated dimensional values can be parametrically adjusted such that they are increasingly integrated to gradually increase the demand on EF. This kind of approach has already been undertaken with respect to the DCCS (Beck et al., 2011; Carlson & Zelazo, 2014), and we suggest that similar efforts to enhance the sensitivity and range of other child EF measures have the potential to advance the study of EF in childhood insofar as they can facilitate a range of research goals. We discuss in more detail two specific examples of established child EF tasks that could be modified in the aforementioned ways in order to further illustrate how altering task features could enhance or reduce difficulty to increase measurement sensitivity.

In the Day/Night Stroop (Gerstadt, Hong, & Diamond, 1994), children are shown pictures of a sun and moon and are instructed to say “night” to the sun and “day” to the moon. This tends to be difficult for preschool-aged children because of pre-established associations between “night” and “moon”, and “day” and “sun”. Our findings suggest this task could be made more challenging by adding several trials in which the usual associations are practiced before switching to the “opposite” rules. Conversely, the task could be made less challenging by reducing the salience of the stimuli (e.g., removing colors and using less iconic depictions of the sun and moon), or by emphasizing conflict in the statement of the rules (e.g., “When I show you a picture of the sun, you do *not* say day. *No way*! You say night.”).

In the Hand Game (Hughes, 1998; Luria, Pribram, & Homskaya, 1964), children are instructed to imitate gestures made by an experimenter (e.g., a fist or a point) for five trials and then are instructed to instead do the opposite of what the experimenter does (e.g., make a fist when the experimenter makes a point). To make this task easier, one might introduce the anti-imitation phase by using contrastive language (e.g., “Now we’re going to play a *different* game where you do not make the same shape as me. *No way!* You make a *different* shape…”), or introduce each trial by prompting children to name the shape they are to make in response to the experimenter’s gesture (e.g., “What shape do you make?”). To make the task more difficult, the number of imitation trials could be increased, which would be expected to enhance the prepotency of the conflicting rules during the anti-imitation phase. These are just a few ways the current findings suggest these tasks could be modified to titrate demands on EF and increase the range and sensitivity of child EF measures.

Limitations of the current meta-analysis

While meta-analyses can be very informative, they also have inherent limitations. We highlight some of these limitations with respect to the current study. First, despite our efforts to identify and control for likely covariates in our analyses, it is possible that unknown and unmeasured variables explain systematic variance between conditions and account for some of the patterns of findings reported here. Second, although we were able to reliably code key variables of interest, the coding necessarily generalized over many differences between conditions. For example, modifications to perceptual salience were actualized in a variety of ways. While it is a strength of the meta-analysis that effects were found despite such heterogeneity, null findings may reflect type II error due to insufficient power to detect effects in the presence of substantial procedural and methodological variations. Third, although we made great efforts to locate and include unpublished findings, we cannot know how many unpublished studies were missed. Meta-analyses that fail to include unpublished studies may overestimate the strength of relationships between variables and can be biased toward positive findings (Lipsey & Wilson, 2001). These issues do not undermine the informativeness of meta-analyses, but they highlight the importance of following up meta-analytic findings with carefully designed replication studies where appropriate.

Conclusion

The meta-analysis reported here clarifies and extends our knowledge of the DCCS, and generates theoretical and practical insights that together have the potential to advance research on the development of EF in early childhood. The findings suggest new directions for research on the developmental mechanisms underlying EF and indicate new approaches to the development of sensitive child EF measures that can facilitate a range of research goals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award 5T32HD007151 from the NICHD to the first author. The authors thank Stephanie Carlson, Melissa Koenig, Yuko Munakata, Sashank Varma, and anonymous reviewers for comments on earlier drafts of the manuscript; Brenton McMenamin for assistance with the analysis; David Wilson for the meta-regression macros used to conduct the analysis; and Jerome Hoover and Zach Machacek for research assistance; and all of the researchers who graciously responded to our requests and provided data for our analysis.

Appendix: Supplemental materials

Supplemental materials can be found online at doi: 10.1016/j.dr.2015.09.001.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition. 1994; 50:7–15.10.1016/0010-0277(94)90018-3 [PubMed: 8039375]

*. Beck DM, Schaefer C, Pang K, Carlson SM. Executive function in preschool children: Test–retest reliability. Journal of Cognition and Development. 2011; 12:169–193.10.1080/15248372.2011.563485 [PubMed: 21643523]

*. Bernier A, Carlson SM, Deschênes M, Matte-Gagné C. Social factors in the development of early executive functioning: A closer look at the caregiving environment. Developmental Science. 2012; 15:12–24.10.1111/j.1467-7687.2011.01093.x [PubMed: 22251288]

Biederman J, Monuteaux MC, Doyle AE, Seidman LJ, Wilens TE, Ferrero F, et al. Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. Journal of Consulting and Clinical Psychology. 2004; 72:757–766.10.1037/0022-006X.72.5.757 [PubMed: 15482034]

*. Bierman KL, Nix RL, Greenberg MT, Blair C, Domitrovich CE. Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. Development and Psychopathology. 2008; 20:821–843.10.1017/S0954579408000394 [PubMed: 18606033]

Blair C, Razza RP. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development. 2007; 78:647–663.10.1111/j.1467-8624.2007.01019.x [PubMed: 17381795]

*. Bohlmann NL, Fenson L. The effects of feedback on perseverative errors in preschool aged children. Journal of Cognition and Development. 2005; 6:119–131.10.1207/s15327647jcd0601_7

Borenstein, M.; Hedges, LV.; Higgins, JPT.; Rothstein, HR. Introduction to meta-analysis. West Sussex: Wiley and Sons, Ltd; 2009.

Botvinick MM, Braver TS. Motivation and cognitive control: From behavior to neural mechanism. Annual Review of Psychology. 2015; 66:83–113.10.1146/annurev-psych-010814-015044

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. Conflict monitoring and cognitive control. Psychological Review. 2001; 108:624–652. [PubMed: 11488380]

*. Brace JJ, Morton JB, Munakata Y. When actions speak louder than words: Improving children's flexibility in a card-sorting task. Psychological Science. 2006; 17:665–669.10.1111/j.1467-9280.2006.01763.x [PubMed: 16913947]

*. Brooks PJ, Hanauer JB, Padowska B, Rosman H. The role of selective attention in preschoolers' rule use in a novel dimensional card sort. Cognitive Development. 2003a; 18:195–215.10.1016/S0885-2014(03)00020-0

*. Brooks PJ, Hanauer JB, Padowska B, Rosman H. Unpublished raw data. 2003b

*. Buss A, Spencer J. Unpublished raw data. 2012

Carlson SM. Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology. 2005; 28:595–616.10.1207/s15326942dn2802_3 [PubMed: 16144429]

*. Carlson SM, Moses LJ. Individual differences in inhibitory control and children's theory of mind. Child Development. 2001; 72:1032–1053.10.1111/1467-8624.00333 [PubMed: 11480933]

*. Carlson SM, White R. Unpublished raw data. 2011

Carlson, SM.; Zelazo, PD. Minnesota executive function scale: test manual. Reflection Sciences, LLC. St. Paul, MN: LLC; 2014.

*. Chatham CH, Yerys BE, Munakata Y. Why won't you do what I want? The informative failures of children and models. Cognitive Development. 2012; 27:349–366.10.1016/j.cogdev.2012.07.003 [PubMed: 24453404]

Chevalier N, Blaye A. Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. Developmental Psychology. 2009; 45:782.10.1037/a0015409 [PubMed: 19413431]

Cools R, Clark L, Owen AM, Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. The Journal of Neuroscience. 2002; 22:4563–4567. [PubMed: 12040063]

Cunningham WA, Zelazo PD. Attitudes and evaluations: A social cognitive neuroscience perspective. Trends in Cognitive Sciences. 2007; 11:97–104.10.1016/j.tics.2006.12.005 [PubMed: 17276131]

Deák GO, Bauer PJ. The dynamics of preschoolers' categorization choices. Child Development. 1996; 67:740–767.10.1111/j.1467-8624.1996.tb01762.x [PubMed: 8706524]

Diamond A. Executive functions. Annual Review of Psychology. 2013; 64:135–168.10.1146/annurev-psych-113011-143750

*. Diamond A, Carlson SM, Beck DM. Preschool children's performance in task switching on the Dimensional Change Card Sort task: Separating the dimensions aids the ability to switch. Developmental Neuropsychology. 2005; 28:689–729.10.1207/s15326942dn2802_7 [PubMed: 16144433]

*. Dick AS, Overton WF, Kovacs SL. The development of symbolic coordination: Representation of imagined objects, executive function, and theory of mind. Journal of Cognition and Development. 2005; 6:133–161.10.1207/s15327647jcd0601_8

Doebel, S.; Zelazo, PD. The role of conceptual knowledge in the development of executive function in early childhood. Paper presented at the 2013 Biennial Meeting of the Cognitive Development Society; Memphis, TN. 2013.

Emerson MJ, Miyake A. The role of inner speech in task switching: A dual-task investigation. Journal of Memory and Language. 2003; 48:148–168.10.1016/S0749-596X(02)00511-9

*. Espinet SD, Anderson JE, Zelazo PD. N2 amplitude as a neural marker of executive function in young children: An ERP study of children who switch versus perseverate on the Dimensional Change Card Sort. Developmental Cognitive Neuroscience. 2011; 2:S49–S58.10.1016/j.dcn.2011.12.002 [PubMed: 22682910]

*. Farrant BM, Maybery MT, Fletcher J. Language, cognitive flexibility, and explicit false belief understanding: longitudinal analysis in typical development and specific language impairment. Child Development. 2012; 83:223–235.10.1111/j.1467-8624.2011.01681.x [PubMed: 22188484]

Ferguson CJ, Heene M. A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. Perspectives on Psychological Science. 2012; 7:555–561.10.1177/1745691612459059 [PubMed: 26168112]

*. Fisher AV. Automatic shifts of attention in the Dimensional Change Card Sort task: Subtle changes in task materials lead to flexible switching. Journal of Experimental Child Psychology. 2011a; 108:211–219.10.1016/j.jecp.2010.07.001 [PubMed: 20674930]

*. Fisher AV. Processing of perceptual information is more robust than processing of conceptual information in preschool-age children: Evidence from costs of switching. Cognition. 2011b; 119:253–264.10.1016/j.cognition.2011.01.015 [PubMed: 21349507]

*. Fosco W, Montgomery DE. Unpublished raw data. 2012

*. Frye D, Zelazo PD, Palfai T. Theory of mind and rule-based reasoning. Cognitive Development. 1995; 10:483–527.10.1016/0885-2014-(95)90024-1s

Garner, WR. The processing of information and structure. Oxford: Lawrence Erlbaum; 1974.

Garner WR, Felfoldy GL. Integrality of stimulus dimensions in various types of information processing. Cognitive Psychology. 1970; 1:225–241.10.1016/0010-0285(70)90016-2

Gerstadt CL, Hong YJ, Diamond A. The relationship between cognition and action: Performance of children 3.5-to 7-years-old on a Stroop-like day-night test. Cognition. 1994; 53:129–153.10.1016/0010-0277(94)90068-X [PubMed: 7805351]

Grant DA, Berg E. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. Journal of Experimental Psychology. 1948; 38:404–411.10.1037/h0059831 [PubMed: 18874598]

Grant DA, Cost JR. Continuities and discontinuities in conceptual behavior in a card sorting problem. The Journal of General Psychology. 1954; 50:237–244.10.1080/00221309.1954.9710121

Hampshire A, Owen AM. Fractionating attentional control using event-related fMRI. Cerebral Cortex. 2006; 16:1679–1689.10.1093/cercor/bhj116 [PubMed: 16436686]

*. Hanania R. Two types of perseveration in the Dimension Change Card Sort task. Journal of Experimental Child Psychology. 2010; 107:325–336.10.1016/j.jecp.2010.05.002 [PubMed: 20566206]

Hanania R, Smith LB. Selective attention and attention switching: Towards a unified developmental approach. Developmental Science. 2010; 13:622–635.10.1111/j.1467-7687.2009.00921.x [PubMed: 20590726]

*. Henning A, Spinath FM, Aschersleben G. The link between preschoolers' executive function and theory of mind and the role of epistemic states. Journal of Experimental Child Psychology. 2011; 108:513–531.10.1016/j.jecp.2010.10.006 [PubMed: 21118749]

*. Hongwanishkul D, Happaney KR, Lee WS, Zelazo PD. Assessment of hot and cool executive function in young children: Age-related changes and individual differences. Developmental Neuropsychology. 2005; 28:617–644.10.1207/s15326942dn2802_4 [PubMed: 16144430]

*. Honomichl RD, Chen Z. Relations as rules: The role of attention in the Dimensional Change Card Sort task. Developmental Psychology. 2011; 47:50–60.10.1037/a0021025 [PubMed: 20954760]

Hughes C. Executive function in preschoolers: Links with theory of mind and verbal ability. British Journal of Developmental Psychology. 1998; 16:233–253.10.1111/j.2044-835X.1998.tb00921.x

Hull CL, Felsinger JM, Gladstone AI, Yamaguchi HG. A proposed quantification of habit strength. Psychological Review. 1947; 54:237–254.10.1037/h0062631 [PubMed: 20265941]

Hunter, JE.; Schmidt, FL. Methods of meta-analysis: Correcting error and bias in research findings. Sage; 2004.

*. Imada T, Carlson SM, Itakura S. East–West cultural differences in context sensitivity are evident in early childhood. Developmental Science. 2011; 16:198–208.10.1111/desc.12016 [PubMed: 23432830]

*. Jacques S, McKay LA. Labels or attributes: Preswitch rule learning strategies and their effects on post-switch performance on the Dimensional Change Card Sort. 2012 Manuscript submitted for publication.

*. Jacques S, Zelazo PD. Unpublished raw data. 1995

*. Jacques S, Zelazo PD, Kirkham NZ, Semcesen TK. Rule selection versus rule execution in preschoolers: An error-detection approach. Developmental Psychology. 1999; 35:770–780.10.1037/0012-1649.35.3.770 [PubMed: 10380867]

*. Jordan PL, Morton JB. Flankers facilitate 3-year-olds' performance in a card-sorting task. Developmental Psychology. 2008; 44:265–274.10.1037/0012-1649.44.1.265 [PubMed: 18194025]

*. Jordan PL, Morton JB. Perseveration and the status of 3-year-olds' knowledge in a card-sorting task: Evidence from studies involving congruent flankers. Journal of Experimental Child Psychology. 2012; 111:52–64.10.1016/j.jecp.2011.06.011 [PubMed: 21862034]

*. Karakelle S, Ertugrul Z. Developmental relations between theory of mind, language, working memory and executive functions. 2011 Manuscript submitted for publication.

*. Kharitonova M, Chien S, Colunga E, Munakata Y. More than a matter of getting "unstuck": Flexible thinkers use more abstract representations than perseverators. Developmental Science. 2009; 12:662–669.10.1111/j.1467-7687.2008.00799.x [PubMed: 19635091]

*. Kirkham NZ, Cruess L, Diamond A. Helping children apply their knowledge to their behavior on a dimension-switching task. Developmental Science. 2003; 6:449–467.10.1111/1467-7687.00300

Kloo D, Perner J. Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. Child Development. 2003; 74:1823–1839.10.1046/j.1467-8624.2003.00640.x [PubMed: 14669898]

*. Kloo D, Perner J. Disentangling dimensions in the dimensional change card-sorting task. Developmental Science. 2005; 8:44–56.10.1111/j.1467-7687.2005.00392.x [PubMed: 15647066]

*. Kloo D, Perner J, Aichhorn M, Schmidhuber N. Perspective taking and cognitive flexibility in the Dimensional Change Card Sorting (DCCS) task. Cognitive Development. 2010; 25:208–217.10.1016/j.cogdev.2010.06.001

*. Kloo, D.; Perner, J.; Giritzer, T. Object-based set-shifting in preschoolers: relations to theory of mind. In: Sokol, BW.; Allen, J.; Carpendale, JIM.; Iarocci, G.; Müller, U.; Young, A., editors.

Self- and social-regulation: Exploring the relations between social interaction, social cognition, and the development of executive functions. Oxford: Oxford University Press; 2010. p. 193-218.

*. Kloo D, Perner J, Kerschhuber A, Dabernig S, Aichhorn M. Sorting between dimensions: Conditions of cognitive flexibility in preschoolers. Journal of Experimental Child Psychology. 2008; 100:115–134.10.1016/j.jecp.2007.12.003 [PubMed: 18241879]

Kray J, Eber J, Karbach J. Verbal self-instructions in task switching: A compensatory tool for action-control deficits in childhood and old age? Developmental Science. 2008; 11:223–236.10.1111/j.1467-7687.2008.00673.x [PubMed: 18333979]

Kray J, Kipp KH, Karbach J. The development of selective inhibitory control: The influence of verbal labeling. Acta Psychologica. 2009; 130:48–57.10.1016/j.actpsy.2008.10.006 [PubMed: 19084817]

*. Lahat A, Degnan KA, White LK, McDermott JM, Henderson HA, Lejuez CW, et al. Temperamental exuberance and executive function predict propensity for risk taking in childhood. Development and Psychopathology. 2012; 24:847–856.10.1111/j.1467-8624.1996.tb01747.x [PubMed: 22781858]

*. Lang B, Perner J. Understanding of intention and false belief and the development of self-control. British Journal of Developmental Psychology. 2002; 20:67–76.10.1348/026151002166325

Lipsey, MW.; Wilson, DB. Practical meta-analysis. London: Sage; 2001.

Littell, JH.; Corcoran, J.; Pillai, V. Systemic reviews and meta-analysis. New York: Oxford University Press; 2008.

Luria AR, Pribram KH, Homskaya ED. An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). Neuropsychologia. 1964; 2:257–280.10.1016/0028-3932(64)90034-X

*. Mack W. Improving postswitch performance in the dimensional change card-sorting task: The importance of the switch and of pretraining by redescribing the test cards. Journal of Experimental Child Psychology. 2007; 98:243–251.10.1016/j.jecp.2007.05.004 [PubMed: 17599343]

MacLeod CM. Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin. 1991; 109:163–203.10.1037/0033-2909.109.2.163 [PubMed: 2034749]

*. Marcovitch S, Boseovski JJ, Knapp RJ. Use it or lose it: Examining preschoolers' difficulty in maintaining and executing a goal. Developmental Science. 2007; 10:559–564.10.1111/j.1467-7687.2007.00611.x [PubMed: 17683342]

*. Marcovitch S, Boseovski JJ, Knapp RJ, Kane MJ. Goal neglect and working memory capacity in 4- to 6-year-old children. Child Development. 2010; 81:1687–1695.10.1111/j.1467-8624.2010.01503.x [PubMed: 21077857]

Marcovitch S, Zelazo PD. The A-Not-B error: Results from a logistic meta-analysis. Child Development. 1999; 70:1297–1313.10.1111/1467-8624.00095

Marcovitch S, Zelazo PD. A hierarchical competing systems model of the emergence and early development of executive function. Developmental Science. 2009; 12:1–18.10.1111/j.1467-7687.2008.00754.x [PubMed: 19120405]

Mazzocco MM, Kover ST. A longitudinal assessment of executive function skills and their association with math performance. Child Neuropsychology. 2007; 13:18–45.10.1080/09297040600611346 [PubMed: 17364562]

*. Miller MR, Giesbrecht GF, Müller U, McInerney RJ, Kerns KA. A latent variable approach to determining the structure of executive function in preschool children. Journal of Cognition and Development. 2012; 13:395–423.10.1111/1467-8624.00095

Miyake A, Emerson MJ, Padilla F, Ahn JC. Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. Acta Psychologica. 2004; 115:123–142.10.1016/j.actpsy.2003.12.004 [PubMed: 14962397]

Miyake A, Friedman NP. The nature and organization of individual differences in executive functions four general conclusions. Current Directions in Psychological Science. 2012; 21:8–14.10.1177/0963721411429458 [PubMed: 22773897]

Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychology. 2000; 41:49–100.10.1006/cogp.1999.0734 [PubMed: 10945922]

Moffitt TE, Arseneault L, Belsky D, Dickson N, Hancox RJ, Harrington HL, et al. A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the National Academy of Sciences. 2011; 108:2693–2698.10.1073/pnas.1010076108

*. Moriguchi Y, Evans AD, Hiraki K, Itakura S, Lee K. Cultural differences in the development of cognitive shifting: East–West comparison. Journal of Experimental Child Psychology. 2011; 111:156–163.10.1016/j.jecp.2011.09.001 [PubMed: 21967676]

*. Moriguchi Y, Hiraki K. Neural origin of cognitive shifting in young children. Proceedings of the National Academy of Sciences. 2009; 106:6017–6021.10.1073/pnas.0809747106

*. Moriguchi Y, Itakura S. Young children's difficulty with inhibitory control in a social context. Japanese Psychological Research. 2008; 50:87–92.10.1111/j.1467-7687.2009.00860.x

*. Moriguchi Y, Kanda T, Ishiguro H, Itakura S. Children perseverate to a human's actions but not to a robot's actions. Developmental Science. 2010; 13:62–68.10.1111/j.1467-7687.2009.00860.x [PubMed: 20121863]

*. Moriguchi Y, Lee K, Itakura S. Social transmission of disinhibition in young children. Developmental Science. 2007; 10:481–491.10.1111/j.1467-7687.2007.00601.x [PubMed: 17552937]

Morton JB, Munakata Y. Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. Developmental Psychobiology. 2002; 40:255–265.10.1002/dev.10033 [PubMed: 11891637]

*. Müller U, Dick AS, Gela K, Overton WF, Zelazo PD. The role of negative priming in preschoolers' flexible rule use on the Dimensional Change Card Sort task. Child Development. 2006; 77:395–412.10.1111/j.1467-8624.2006.00878.x [PubMed: 16611180]

*. Müller U, Kerns KA, Konkin K. Test–retest reliability and practice effects of executive function tasks in preschool children. The Clinical Neuropsychologist. 2012; 26:271–287.10.1080/13854046.2011.645558 [PubMed: 22256888]

Müller U, Zelazo PD, Hood S, Leone T, Rohrer L. Interference control in a new rule use task: Age-related changes, labeling, and attention. Child Development. 2004; 75:1594–1609.10.1111/j.1467-8624.2004.00759.x [PubMed: 15369533]

*. Müller U, Zelazo PD, Imrisek S. Executive function and children's understanding of false belief: How specific is the relation? Cognitive Development. 2005; 20:173–189.

*. Müller U, Zelazo PD, Lurye LE, Liebermann DP. The effect of labeling on preschool children's performance in the Dimensional Change Card Sort task. Cognitive Development. 2008; 23:395–408.10.1016/j.cogdev.2008.06.001

Munakata Y. Infant perseveration and implications for object permanence theories: A PDP model of the A-not-B task. Developmental Science. 1998; 1:161–211.10.1111/1467-7687.00021

*. Munakata Y, Yerys BE. All together now: When dissociations between knowledge and action disappear. Psychological Science. 2001; 12:335–337.10.1111/1467-9280.00361 [PubMed: 11476102]

*. Nguyen TK, Astington JW. Unpublished raw data. 2011

*. O'Leary A, Marcovitch S. Unpublished raw data. 2011

*. O'Neill G, Miller PH. A show of hands: Relations between young children's gesturing and executive function. Developmental Psychology. 2013; 49:1517–1528.10.1037/a0030241 [PubMed: 23046430]

Pennington BF, Ozonoff S. Executive functions and developmental psychopathology. Journal of Child Psychology and Psychiatry. 1996; 37:51–87.10.1111/j.1469-7610.1996.tb01380.x [PubMed: 8655658]

*. Perner J, Lang B. What causes 3-year-olds' difficulty on the Dimensional Change Card Sorting task? Infant and Child Development. 2002; 11:93–105.10.1111/1467-8624.00436

*. Perner J, Lang B, Kloo D. Theory of mind and self-control: More than a common problem of inhibition. Child Development. 2002; 73:752–767.10.1111/1467-8624.00436 [PubMed: 12038549]

*. Qu L, Zelazo PD. The facilitative effect of positive stimuli on 3-year-olds' flexible rule use. Cognitive Development. 2007; 22:456–473.10.1016/j.cogdev.2007.08.010

*. Ramscar M, Dye M, Witten J, Klein J. Dual routes to cognitive flexibility: Learning and response conflict resolution in the Dimensional Change Card Sort task. Child Development. 2013; 84:1308–1323.10.1111/cdev.12044 [PubMed: 23311677]

*. Rennie DAC, Bull R, Diamond A. Executive functioning in preschoolers: Reducing the inhibitory demands of the Dimensional Change Card Sort task. Developmental Neuropsychology. 2004; 26:423–443.10.1207/s15326942dn2601_4 [PubMed: 15276903]

Richland L, Burchinal M. Early executive function predicts reasoning development. Psychological Science. 2012; 24:87–92.10.1177/0956797612450883 [PubMed: 23184588]

Rosenthal R. The file drawer problem and tolerance for null results. Psychological Bulletin. 1979; 86:638–641.10.1037/0033-2909.86.3.638

Schooler J. Unpublished results hide the decline effect. Nature. 2011; 470:437.10.1038/470437a [PubMed: 21350443]

Smith LB, Kemler DG. Levels of experienced dimensionality in children and adults. Cognitive Psychology. 1978; 10:502–532.10.1016/0010-0285(78)90009-9 [PubMed: 699515]

*. Towse JN, Redbond J, Houston-Price CM, Cook S. Understanding the Dimensional Change Card Sort: Perspectives from task success and failure. Cognitive Development. 2000; 15:347–365.10.1016/S0885-2014(00)00021-6

*. van Bers BMCW, Visser I, van Schijndel TJP, Mandell DJ, Raijmakers MEJ. The dynamics of development on the Dimensional Change Card Sorting task. Developmental Science. 2011a; 14:960–971.10.1111/j.1467-7687.2011.01045.x [PubMed: 21884312]

*. van Bers BMCW, Visser I, van Schijndel TJP, Mandell DJ, Raijmakers MEJ. Unpublished raw data. 2011b

*. Vendetti C, Kamawar D. Unpublished raw data. 2011

Vygotsky, LS. Thought and language. Hanfmann, E.; Vakar, G., translators. Cambridge, MA: MIT Press; 1935/1962.

Wiebe SA, Sheffield T, Nelson JM, Clark CA, Chevalier N, Espy KA. The structure of executive function in 3-year-olds. Journal of Experimental Child Psychology. 2011; 108:436–452.10.1016/j.jecp.2010.08.008 [PubMed: 20884004]

Willoughby MT, Blair CB, Wirth RJ, Greenberg M. The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. Psychological Assessment. 2012; 24:226.10.1037/a0025361 [PubMed: 21966934]

*. Wimmer MC, Doherty MJ. The development of ambiguous figure perception. Monographs of the Society for Research in Child Development. 2011; 76:1–130.10.1111/j.1540-5834.2011.00589.x

Winsler, A.; Fernyhough, C.; Montero, I., editors. Private speech, executive functioning, and the development of verbal self-regulation. New York: Cambridge University Press; 2009.

*. Yerys BE, Munakata Y. When labels hurt but novelty helps: Children's perseveration and flexibility in a card-sorting task. Child Development. 2006; 77:1589–1607.10.1111/j.1467-8624.2006.00961.x [PubMed: 17107448]

Zelazo PD. The development of conscious control in childhood. Trends in Cognitive Sciences. 2004; 8:12–17.10.1016/j.tics.2003.11.001 [PubMed: 14697398]

Zelazo PD. The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. Nature Protocols. 2006; 1:297–301.10.1038/nprot.2006.46 [PubMed: 17406248]

Zelazo, PD.; Anderson, JE.; Richler, J.; Wallner-Allen, K.; Beaumont, JL.; Weintraub, S. NIH Toolbox Cognition Battery (NIHTB-CB): Measuring executive function and attention. In: Zelazo, PD.; Bauer, PJ., editors. National Institutes of Health Toolbox Cognition Battery (NIHTB-CB): Validation for children between 3 and 15 years. Vol. 78. Monographs of the Society for Research in Child Development; 2013. p. 1-172.

Zelazo PD, Carlson SM. Hot and cool executive function in childhood and adolescence: Development and plasticity. Child Development Perspectives. 2012; 6:354–360.10.1111/j.1750-8606.2012.00246.x

Zelazo, PD.; Doebel, S. The role of reflection in promoting adolescent self-regulation. In: Oettingen, G.; Gollwitzer, P., editors. Self-regulation in adolescence. Cambridge University Press; 2015. p. 212-239.

*. Zelazo PD, Frye D, Rapus T. An age-related dissociation between knowing rules and using them. Cognitive Development. 1996; 11:37–63.10.1016/S0885-2014(96)90027-1

*. Zelazo PD, Müller U, Frye D, Marcovitch S. The development of executive function in early childhood. Monographs of the Society for Research in Child Development. 2003; 68:vii–137.10.1111/j.1540-5834.2003.06803001.x [PubMed: 14723273]

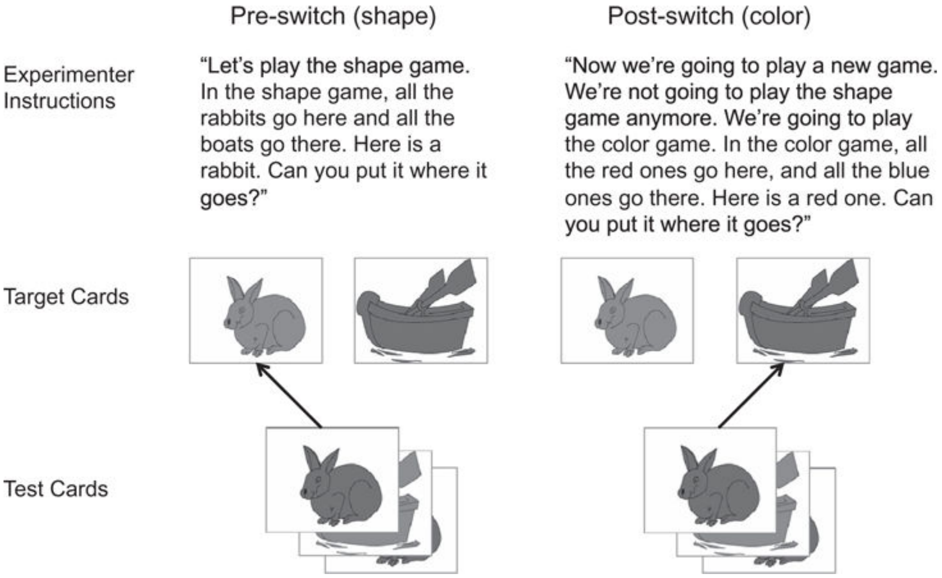

Fig. 1.

Sample stimuli and instructions used in the standard version of the DCCS, based on Zelazo (2006). Dark gray represents blue and light gray represents red.

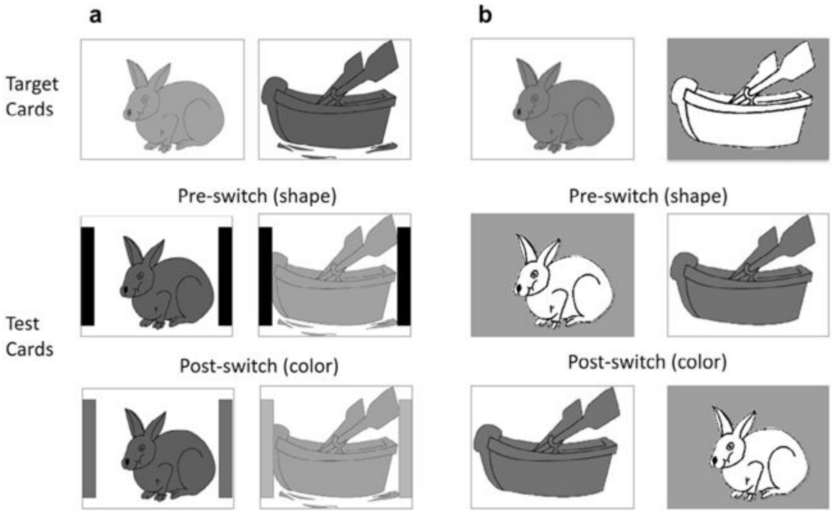

Fig. 2.

Examples of experimental modifications to the DCCS. (a) DCCS task stimuli that enhanced the salience of the post-switch phase, modeled on Jordan and Morton (2008), with black (neutral) bars flanking the pre-switch test cards during the shape game, and relevant, colored bars flanking the post-switch test cards during the color game. (b) Spatially separated stimuli, modeled on Diamond et al. (2005). White represents red, and gray represents blue.

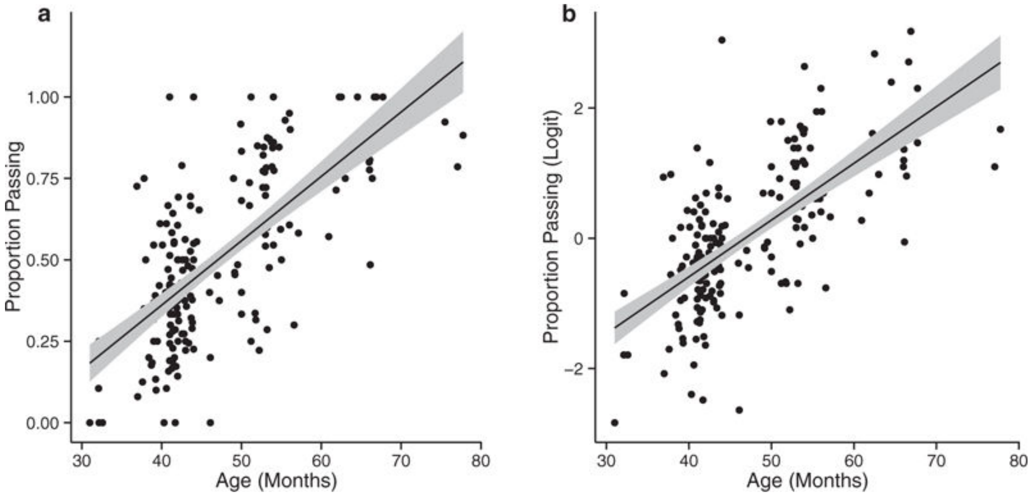

Fig. 3.

Proportion passing post-switch phase as a function of age: (a) raw proportion and (b) logit transformed. Shaded regions indicate the 95% confidence interval.

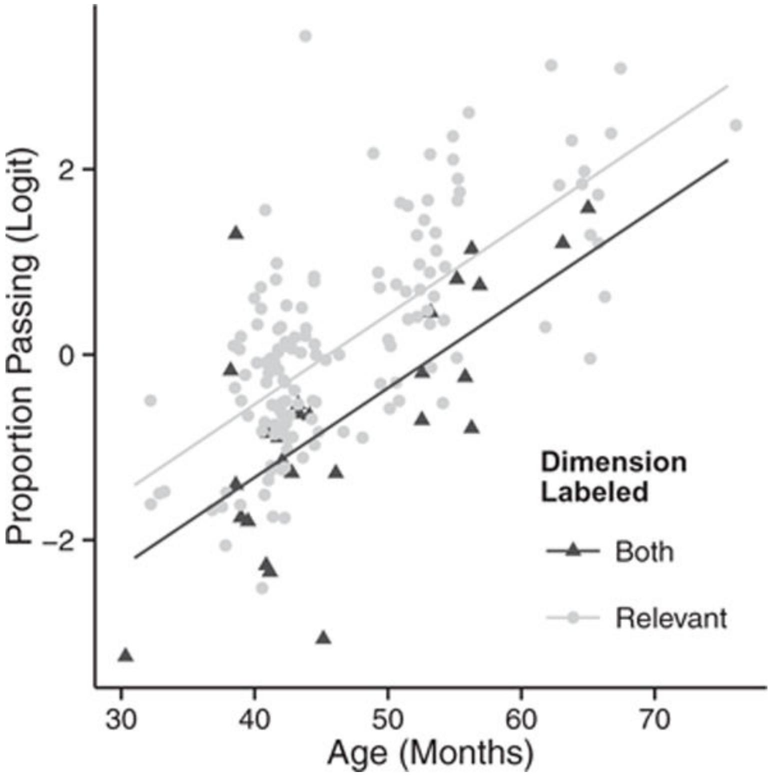

Fig. 4.

Relation between switching (proportion passing) and age by labeling procedure. These data suggest that, across the age groups studied, children tend to switch more when only the relevant dimension is labeled by the experimenter versus when both dimensions are labeled.

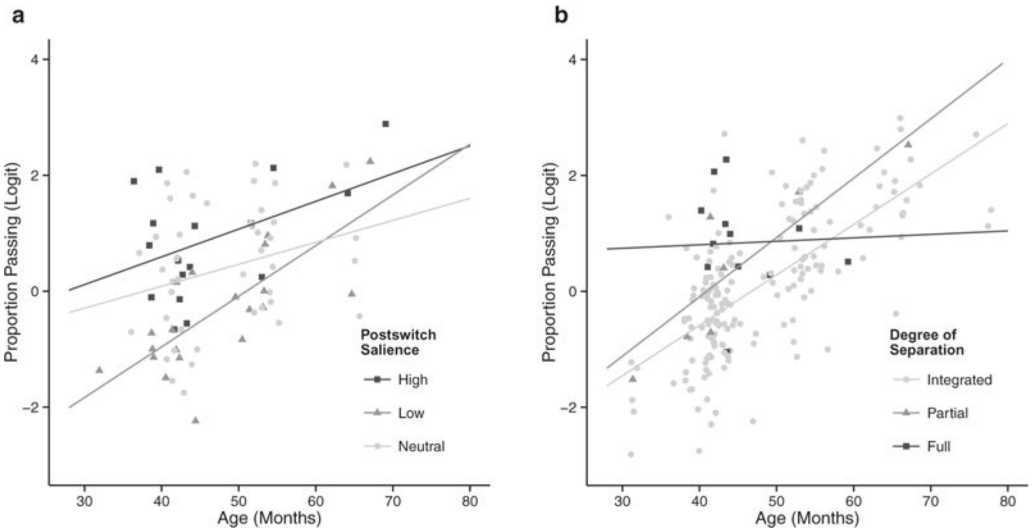

Fig. 5.

Relation between switching (proportion passing) and age by (a) relative post-switch salience and (b) degree of separation.

Table 1

Reports and conditions included in the meta-analysis.

Authors	Year	*k*		Total
		Standard	Experimental	
Beck, Schaefer, Pang, and Carlson	2011	2	4	6
Bernier, Carlson, Deschênes, and Matte-Gagné	2012	1	0	1
Bierman, Nix, Greenberg, Blair, and Domitrovich	2008	3	0	3
Bohlmann and Fenson	2005	3	3	6
Brace, Morton, and Munakata	2006	1	2	3
Brooks, Hanauer, Padowska, and Rosman	2003a	0	10	10
Brooks, Hanauer, Padowska, and Rosman*	2003b	1	1	2
Buss and Spencer*	2012	3	4	7
Carlson and Moses	2001	2	0	2
Carlson and White*	2011	2	0	2
Chatham, Yerys, and Munakata	2012	0	1	1
Diamond, Carlson, and Beck	2005	3	3	6
Dick, Overton, and Kovacs	2005	3	0	3
Doebel and Zelazo*	2012	8	1	9
Espinet, Anderson, and Zelazo	2011	2	0	2
Farrant, Maybery, and Fletcher	2012	3	0	3
Fisher	2011a	0	4	4
Fisher	2011b	0	18	18
Fosco and Montgomery*	2012	3	0	3
Frye, Zelazo, and Palfai	1995	7	0	7
Hanania	2010	4	0	4
Henning, Spinath, and Aschersleben	2011	4	0	4
Hongwanishkul, Happaney, Lee, and Zelazo	2005	3	0	3
Honomichl and Chen	2011	4	14	18
Imada, Carlson, and Itakura*	2011	6	0	6
Jacques and Zelazo*	1995	3	6	9
Jacques and McKay*	2012	2	18	20
Jacques, Zelazo, Kirkham, and Semcesen	1999	3	6	9
Jordan and Morton	2008	2	3	5
Jordan and Morton	2012	0	4	4
Karakelle and Ertugrul*	2011	3	0	3
Kharitonova, Chien, Colunga, and Munakata	2009	2	0	2
Kirkham, Cruess, and Diamond	2003	2	6	8
Kloo and Perner	2003	10	2	12
Kloo and Perner	2005	6	10	16
Kloo, Perner, Aichhorn, and Schmidhuber	2010	1	2	3

Authors	**Year**	***k***		**Total**
		Standard	**Experimental**	
Kloo, Perner, and Giritzer*	2010	2	0	2
Kloo, Perner, Kerschhuber, Dabernig, and Aichhorn	2008	10	18	28
Lahat et al.	2012	2	0	2
Lang and Perner	2002	3	0	3
Mack	2007	2	2	4
Marcovitch, Boseovski, and Knapp	2007	0	8	8
Marcovitch, Boseovski, Knapp, and Kane	2010	0	4	4
Miller, Giesbrecht, Müller, McInerney, and Kerns	2012	3	0	3
Moriguchi and Itakura	2008	2	2	4
Moriguchi, Evans, Hiraki, Itakura, and Lee	2011	4	4	8
Moriguchi, Kanda, Ishiguro, and Itakura	2010	0	2	2
Moriguchi, Lee, and Itakura	2007	0	13	13
Müller, Dick, Gela, Overton, and Zelazo	2006	0	3	3
Müller, Kerns, and Konkin	2012	3	0	3
Müller, Zelazo, and Imrisek	2005	3	0	3
Müller, Zelazo, Lurye, and Liebermann	2008	9	7	16
Munakata and Yerys	2001	1	0	1
Nguyen and Astington*	2011	2	0	2
O'Leary and Marcovitch*	2011	0	9	9
O'Neill and Miller	2013	4	4	8
Perner and Lang	2002	2	6	8
Perner, Lang, and Kloo	2002	4	0	4
Qu and Zelazo	2007	5	5	10
Ramscar, Dye, Witten, and Klein	2013	2	0	2
Rennie, Bull, and Diamond	2004	1	2	3
Towse, Redbond, Houston-Price, and Cook	2000	3	2	5
van Bers, Visser, van Schijndel, Mandell, and Raijmakers	2011a	3	0	3
van Bers, Visser, van Schijndel, Mandell, and Raijmakers*	2011b	0	3	3
Vendetti and Kamawar*	2011	4	4	8
Wimmer and Doherty	2011	3	0	3
Yerys and Munakata	2006	0	2	2
Zelazo, Frye, and Rapus	1996	5	0	5
Zelazo, Müller, Frye, and Marcovitch	2003	6	14	20
Totals		190	236	426

k = number of conditions.

[*] Data unpublished at the time of retrieval.

Table 2

Summary of inverse-variance-weighted multiple regression analyses.

Conditions	Predictor	*B*	*SE*	95% CI	*z*	*p*	β	ES (odds ratio)[a]
Standard	Year of publication	0.01	0.01	[−0.02, 0.04]	0.89	.37	.07	
	Task modality	−0.03	0.13	[−.029, 0.23]	−0.23	.82	−.01	
	Pass criterion	−0.61	2.05	[−4.63, 3.41]	−0.30	.77	−.02	
	Age	0.10	0.08	[0.08, 0.11]	12.26	<.001	.74	3.21
	Labeling procedure	0.38	0.07	[0.24, 0.51]	5.50	<.001	.32	Relevant vs. both dim. labeled = 2.13
	Conflict emphasis	0.19	0.09	[0.03, 0.36]	2.29	.02	.15	Conflict emphasis vs. standard = 1.47
	n pre-switch trials	−0.16	0.07	[−0.29, −0.02]	−2.22	.03	−.15	5 vs. 6 trials = 1.17
	Pre-switch feedback	−0.07	0.07	[−0.20, 0.06]	−1.02	.31	−.07	
	Pre-test practice	−0.18	0.11	[−0.40, 0.05]	−1.56	.19	−.11	
Standard + salience	Age	0.07	0.01	[0.06, 0.08]	11.73	<.001	.57	
	Relative post-switch salience	0.59	0.14	[0.32, 0.86]	4.24	<.001	.21	1.80
Standard + separated	Age	0.08	0.01	[0.06, 0.07]	12.52	<.001	.64	
	Degree of spatial separation	0.46	0.12	[0.15, 0.58]	3.85	<.001	.29	1.59
Standard + alternate[b]	Age	0.07	0.01	[0.05, 0.07]	12.02	<.001	.55	
	Modality	−0.14	0.07	[−0.28, 0.01]	−2.02	.05	−.10	1.33
	Reversal shift – bidimensional	−0.49	0.26	[−1.01, 0.03]	−1.86	.06	−.10	0.61
	Observation	0.11	0.15	[−0.19, 0.41]	0.75	.46	.04	
	Redundant/goal neglect	−0.44	0.23	[−0.88, 0.01]	−1.96	.05	−.11	0.64
	Child label	0.47	0.26	[−0.03, 0.97]	1.83	.07	.10	1.60
	Child label (+labeling procedure)	0.41	0.24	[−0.06, 0.87]	4.51	.09	.24	
	Reversal shift – unidimensional	1.46	0.24	[0.99, 1.93]	6.06	<.001	.31	4.31
	Target character	1.73	0.24	[1.26, 2.20]	7.24	<.001	.37	5.66

CI = confidence interval.

*B*s for all control variables are reported for the first analysis and in subsequent analyses only if significant.

[a] Effect sizes reflect increased odds of switching compared to standard condition given an increase of one level of the independent variable unless otherwise specified.

[b] *B*s for significant control predictors correspond to omnibus analysis; *B*s for each version correspond to separate regression analyses. Effect sizes are only reported for predictors significant at $p < .05$.

Table 3

Raw and adjusted logits and mean proportions for significant predictors in multiple regression analyses.

Conditions	Variable	*k*	Logit				Proportion switching	
			M	*SE*	Adjusted *M*	95% CI	*M*	Adjusted *M*
Standard	Age[a]							
	3 years	79	−0.39	0.06	−0.35	[−0.51, −0.27]	.40	.41
	4 years	29	0.66	0.11	0.82	[0.44, 0.88]	.66	.69
	5 years	10	1.55	0.21	1.99	[1.14, 1.97]	.83	.88
	Labeling procedure							
	Both dimensions labeled	21	−0.44	0.12	−0.61	[−0.68, −0.21]	.39	.35
	Relevant only	100	0.03	0.06	0.15	[−0.09, 0.14]	.51	.54
	Conflict emphasis							
	Standard	103	−0.06	0.06	−0.47	[−0.17, 0.05]	.48	.49
	Heightened	18	−0.06	0.14	−0.05	[−0.32, 0.21]	.49	.58
	n pre-switch trials							
	6 or more trials	65	0.14	0.07	−0.03	[0.00, 0.28]	.54	.49
	Less than 6 trials	56	0.07	0.08	0.13	[−0.09, 0.23]	.52	.53
Standard and salience	Low post-switch salience	20	−0.34	0.19	−0.48	[−0.72, 0.03]	.42	.38
	Neutral	237	0.05	0.05	0.13	[−0.05, 0.15]	.51	.53
	High post-switch salience	17	0.71	0.20	0.72	[0.32, 1.10]	.67	.67
Standard and separated	Integrated	190	0.03	0.05	−0.29	[−0.07, 0.13]	.51	.43
	Partial separation	9	0.16	0.24	0.17	[−0.30, 0.63]	.54	.54
	Full separation	12	0.75	0.26	0.64	[0.23, 1.27]	.68	.65

k = number of conditions; CI = confidence interval.

[a] Age groups with >5 conditions are presented.

Covariate adjusted means (proportions) were computed using full weighted regression models, controlling for age and modality by setting values of these variables to their mean. All independent variables in the models were weighted by the inverse variance and random effects variance component.

Adjusted means computed for midpoint of each age group and for 5 vs. 6 pre-switch trials. Only control variables that were significant in the multiple regression models are displayed.

Table 4

Raw and adjusted logits and mean proportions DCCS versions included in multiple regression analyses.

DCCS version	***k***	***M* age^a (months)**	**Logit**				**Proportion switching**	
			M	***SE***	**95% CI**	**Adjusted *M***	***M***	**Adjusted *M***
Child label$^\wedge$	9	42.39	0.10	0.25	[−0.39, 0.59]	0.48	0.53	0.62
Redundant/goal neglect*	12	57.92	0.35	0.21	[−0.06, 0.76]	−0.36	0.59	0.41
Observation	27	48.26	0.24	0.14	[−0.04, 0.52]	0.16	0.56	0.54
Target character*	13	45.81	1.77	0.22	[1.33, 2.20]	1.76	0.85	0.85
Reversal – unidimensional*	14	42.64	1.19	0.23	[0.74, 1.64]	1.47	0.77	0.81
Reversal – bidimensional$^\wedge$	12	46.30	−0.56	0.26	[−1.07, −0.06]	−0.46	0.36	0.39

k = number of conditions; CI = confidence interval.

*Predicted switching at $p < .05$;

$^\wedge$ $p < .10$.

^{a}Mean age of children in each condition.

Mean age and effect size (logit and proportion) are weighted by the inverse variance and a random effects variance component derived for each model. Effect sizes are converted from logits back to proportions for interpretation. See supplemental materials for descriptions of these versions.

Table 5

Descriptive statistics for alternate versions not included in the regression analyses.

DCCS version	*k*	*M* age (months)	*M* logit	*SE*	*M* proportion switching	95% CI
Hot	6	42.06	0.02	0.17	0.50	[−0.32, 0.36]
Negative priming	6	44.72	0.00	0.19	0.50	[−0.38, 0.37]
No target	6	40.62	0.36	0.23	0.59	[−0.10, 0.82]
Opposite label	3	52.44	0.04	0.23	0.51	[−0.41, 0.49]
Partial	5	41.81	−0.26	0.21	0.44	[−0.67, 0.16]
Partial–partial change	3	37.26	0.00	0.28	0.50	[−0.55, 0.54]
Post-switch feedback	4	39.18	1.65	0.39	0.84	[0.88, 2.41]
Pruned tree	4	53.26	−0.05	0.18	0.49	[−0.41, 0.31]
Redundant pre-switch	4	42.69	0.03	0.27	0.51	[−0.51, 0.57]
Total change	3	40.35	0.98	0.28	0.73	[0.42, 1.54]
Unidimensional	3	42.49	1.43	0.46	0.81	[0.53, 2.34]

k = number of conditions; CI = confidence interval.

Mean age and effect size (logit and proportion) are weighted by the inverse variance. Effect sizes are converted from logits back to proportions for interpretation. See supplemental materials for descriptions of these versions.

Table 6

Summary of influence on difficulty of modifications to the DCCS.

Easier	More difficult	No detected effect
Label relevant dimension Emphasize conflict between sorting games when introducing post-switch phase Reduce number of pre-switch trials Increase salience of post-switch dimension by altering perceptual properties of stimuli Spatially separate dimensional values on test and/or target cards Use cards that vary on only one dimension with reversal instructions Have child label test card instead of experimenter Use target characters instead of target cards Remove or omit use of target cards	Label both dimensions Increase number of pre-switch trials Reduce salience of post-switch dimension by altering perceptual properties of stimuli Include test cards that fully match target in post-switch phase Replace standard instructions with reversal instructions	Provide pre-switch feedback Provide pre-test practice Administer with cards (table top) or on computer Have child observe another individual sort cards during pre-switch phase Have child make judgments of own or other's performance, or point to boxes instead of manually sorting cards