Journal of Experimental Child Psychology 221 (2022) 105428

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Exploring the development of attentional set shifting in young children with a novel Intradimensional/Extradimensional shift task

Eva Reindl [a,b,*], Christoph Johannes Völter [a,c], Jessica Campbell-May [a], Josep Call [a], Amanda Madeleine Seed [a]

[a] School of Psychology and Neuroscience, University of St Andrews, St Andrews KY16 9JP, UK
[b] Department of Anthropology, Durham University, Durham DH1 3LE, UK
[c] Messerli Research Institute, University of Veterinary Medicine Vienna, 1210 Vienna, Austria

A R T I C L E I N F O

Article history:
Received 8 July 2021
Revised 21 January 2022
Available online 27 April 2022

Keywords:
Cognitive flexibility
Attentional set shifting
Attention shifting
Executive functions
Intradimensional/Extradimensional shift
Cognitive development

A B S T R A C T

Attentional set shifting is a core part of cognition, allowing quick and flexible adaption to new demands. The study of its development during early childhood has been hampered by a shortage of measures not requiring language. This article argues for a revival of the Intradimensional/Extradimensional (ID/ED) shift task by presenting a new nonverbal version of the task (Shifting Tray task). Children (*N* = 95 3- to 5-year-olds; 49 girls; predominantly European White) were presented with pairs of trays, each filled with a substrate and an upside-down cup on top, and were asked to find stickers. In the pre-switch phase, children learned (through trial and error) which dimension (substrate or cup) was predictive of the rewards. In the post-switch phase, all stimuli were exchanged. For children in the intradimensional shift condition, the dimension predictive of the sticker was the same as the one predictive in the pre-switch phase. For children in the extradimensional shift condition, the previously irrelevant dimension was now relevant. Results showed that most 3-year-olds were able to switch, and older children did not outperform younger children. The easy and flexible nature of the task allows researchers to investigate the impact of labels and instructions and to use it in cross-cultural and comparative research.

* Corresponding author at: Department of Anthropology, Durham University, Durham DH1 3LE, UK.
E-mail address: eva.reindl@live.de (E. Reindl).

https://doi.org/10.1016/j.jecp.2022.105428
0022-0965/© 2022 The Authors. Published by Elsevier Inc.

Introduction

Executive functions (EFs) encompass a host of cognitive mechanisms allowing for goal-directed behavior, planning, mental flexibility, and self-monitoring (Henry & Bettenay, 2010) and are a core part of cognition. EFs are crucial for regulating other cognitive abilities such as reasoning, problem solving, reading, and writing and are predictive of school readiness, academic, professional, and relationship success, and health behaviors (Best, Miller, & Naglieri, 2011; Blair & Razza, 2007; Bull, Espy, & Senn, 2004; Bull, Espy, & Wiebe, 2008; Diamond, 2013, 2016; St Clair-Thompson & Gathercole, 2006). One crucial component of EFs is the ability to switch attention between rules or mental "sets", allowing quick and flexible adaption to new demands, problems, or instructions (Brown & Tait, 2015a, 2015b; Cragg & Chevalier, 2012; Garon, Bryson, & Smith, 2008; Ionescu, 2012).

Despite the increased research effort during the past decades on attentional control and shifting, as well as on their relation to other EFs, the mechanisms underpinning the development of these abilities are still unclear (Best & Miller, 2010; Burgoyne & Engle, 2020; Cragg & Chevalier, 2012; Doebel & Zelazo, 2015; Fisher, 2019; Garon et al., 2008; Karr et al., 2018). This is partly due to a shortage of tasks suitable for use with young children. For example, although the number of latent variable studies investigating the structure of EFs in preschool children has increased over the past few years, many have been criticized for not including measures of attentional set shifting or doing so with only a single task (Bardikoff & Sabbagh, 2017; Karr et al., 2018). Additionally, as shown below, many of the existing tasks rely heavily on children's language comprehension skills and understanding of arbitrary rules. This arguably adds unnecessary additional cognitive load (which makes some tasks not suitable for use with children under 4 years of age) and makes the tasks even more "impure" measures of attentional set shifting (Burgess, 1997; Miyake et al., 2000). Therefore, the first goal of the current study was to revive a measure of attentional set shifting that does not rely on verbal instructions—the Intradimensional/Extradimensional (ID/ED) attentional set shifting task (Brown & Tait, 2015a; Esposito, 1975; Roberts, Robbins, & Everitt, 1988; Slamecka, 1968; Wolff, 1967)—by presenting a new version of the task that can be used with children as young as 3 years. A second goal of the current study was to investigate potential age differences in performance across the preschool years.

Tasks assessing attentional set shifting in preschool children

Currently, the most widely used measure for attentional set shifting in children is the Dimensional Change Card Sort (DCCS) task (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006), in which children are asked to match test cards displaying bidimensional stimuli (e.g., a blue boat or a red rabbit) to one of two target cards (e.g., a red boat and a blue rabbit) according to one dimension (e.g., color) in the pre-switch phase and according to the other dimension (e.g., shape) in the post-switch phase. In the standard version of the DCCS, the majority of 3-year-olds fail to shift attention to the new rule in the post-switch phase, whereas most 5-year-olds switch successfully (Doebel & Zelazo, 2015).

There are several other attentional set shifting tasks for young children—most of which also make use of explicit verbal instructions to direct attention. Similar to the DCCS, the Preschool Attentional Switching Task (PAST; Chevalier & Blaye, 2008), the Flexible Induction of Meaning (FIM) task (Deák, 2000), the Trail Making Test for Preschoolers (TRAILS-P; Espy & Cwik, 2004), the Shape School (Espy, 1997), the Object Classification Task for Children (OCTC; Smidts, Jacobs, & Anderson, 2004), and the Switching Inhibition and Flexibility Task (SwIFT; FitzGibbon, Cragg, & Carroll, 2014) all require children to respond flexibly to pictures or objects, as well as to shift attention, in response to an explicit instruction from an experimenter to use a particular rule or stimulus feature. All tasks show marked improvement over the 3- to 5-year age range. An important source of this developmental improvement might involve an increased ability to use labels to direct attention (Buss & Nikam, 2020; Hanania & Smith, 2010). In all these tasks, verbal labels are provided in the experimenter's ongoing instructions. Therefore, the extent to which preschoolers could use their increasing knowledge of categories and labels to shift their attention without such explicit instruction becomes not clear from these studies.

One task that has been used to examine attentional set shifting without an explicit instruction to attend to a particular dimension is the Triad Classification (TC) task (Smith & Kemler, 1977). This object matching task was originally developed to assess selective attention to dimensions. Children are presented with a picture of an object with two features (color and shape) and a choice of two others to match with it. Older children, who attend more selectively toward a single dimension, tend to prefer objects that match precisely on one dimension (e.g., same color), whereas younger children seemingly process information more holistically and tend to choose objects that are more similar when multiple dimensions are considered despite lacking an exact match on any one dimension. A recent study examined the relation between 3- and 4-year-old's performance in the TC and the DCCS tasks and found that children who successfully switched attention in the post-switch phase of the DCCS were more likely to match on one dimension and also to successfully switch between dimensions in the TC task, demonstrating a greater ability to selectively attend to dimensions across tasks (Buss & Kerr-German, 2019). The study showed a correlation between a task in which the relevant dimension is explicitly labeled (DCCS) and one in which the instruction is more abstract (i.e., to match cards; TC). Nevertheless, as in the other tasks cited above, the TC task requires children to complete a verbal request to match or sort according to a rule.

Explaining the development of attentional set shifting during the preschool years

Some theories suggest a conceptual change at the core of the developmental trend in attention shifting. The *cognitive complexity and control theory* (Zelazo et al., 2003) proposes that switching in the DCCS requires an integration of the if–then sorting rules into a hierarchy of setting conditions (i.e., "under which condition [shape game or color game] does which set of if–then rules apply?"), that is, an understanding of if–if–then structures. Three-year-olds are assumed to be not yet capable of representing these complex rule hierarchies, although they are able to hold several if–then rules in mind at a time. According to the *representational redescription account* (Kloo & Perner, 2005; Perner & Lang, 2002), 3-year-olds are assumed to not yet be able to represent an object (here the image on the test card) along more than one dimension, preventing them from switching to the new sorting rule.

Other theories place more emphasis on domain-general cognitive processes. The *attentional inertia theory* (Kirkham et al., 2003) suggests that although 3-year-olds are able to focus their attention on the level of a dimension (e.g., color or shape), their inhibition capacities are still too weak to allow them to overcome this focus on what was previously relevant when the rules change (Landry, Al-Taie, & Franklin, 2017). However, proponents of other theories have noted that this account is still somewhat underspecified, for example, lacking a precise definition of attention or a detailed description of the processes underlying perseveration (Buss & Spencer, 2014; Munakata, Morton, & Yerys, 2003).

The *dynamic neural field theory* aims to move beyond the mere descriptive account of the attentional inertia theory and proposes a neurally grounded model to more precisely describe the mechanisms underlying selective attention and attention shifting as well as the developmental changes (Buss & Kerr-German, 2019; Buss & Nikam, 2020; Buss & Spencer, 2014). It assumes the presence of a dimensional attentional system in the frontal cortical areas, capturing the representation of dimensional labels such as "color" and "shape". In this system, neural units interact through a local-excitation/lateral-inhibition function, resulting in localized peaks of activation that reflect the representation of or decision for a dimension (Buss & Kerr-German, 2019; Buss & Spencer, 2018). The frontal dimensional attentional system is connected to neural systems in the temporal and parietal lobes that encode the specific stimulus features such as shape (e.g., circle, square), color (e.g., yellow, blue), and the spatial location of the stimuli in the environment (see Fig. 3 in Buss & Kerr-German, 2019, or Fig. 1 in Buss & Spencer, 2018). The link between the frontal and posterior neural areas is bidirectional. The model proposes that older children's more categorical, all-or-none attention toward dimensions stems from stronger neural connections both within the frontal area and between the frontal and posterior areas. Thus, when prompted with the dimension label "shape", activity in the frontal dimensional attention system can more effectively activate the dimension shape and inhibit the dimension color and can more effectively boost activation in the associated posterior areas, resulting in more enhanced processing of the shape features of the stimuli (top-down influence). In addition, repeated sorting along one dimension more strongly activates the corresponding dimension label in the frontal area (bottom-up influence) (Buss & Spencer, 2018).

The *selective attention theory* also emphasizes the importance of attending to the level of dimension for success on the DCCS and that over the preschool years capacity for selective attention toward dimensions per se increases (Benitez, Vales, Hanania, & Smith, 2017; Hanania, 2010; Hanania & Smith, 2010). Selective attention describes the enhanced processing of some parts of the sensory input while dismissing others, in other words, a differential weighting of input (Fisher, 2019; Hanania & Smith, 2010). Selective attention toward one dimension (and dismissal of another dimension) can be visualized to lie on a continuum, with one end being a very categorical, all-or-none-like division of attention and the other end representing a more graded distribution of attention (see Fig. 7 in Hanania & Smith, 2010). For example, in the DCCS, when sorting by color, focusing on features would mean learning about the specific color values (e.g., red, blue) but not about color rules in general. Hanania (2010) found evidence that some children indeed focus on features rather than dimensions in the DCCS. In that study, a third phase was administered after the post-switch phase of the DCCS, in which 3- and 4-year-olds needed to sort along the same dimension as in the post-switch phase, but now the cards had new colors and shapes. Whereas 64% of the children who failed the post-switch phase still perseverated on the wrong *dimension* in this third stage (as predicted by the attentional inertia theory), one third of children now passed, indicating that these children had been fixated only on the specific *features* of the color/shape but not on the dimension per se.

Both the selective attention theory and dynamic neural field theory agree that successful switching in the DCCS in older preschoolers is facilitated by an ability to apply attention in a discrete or all-or-none fashion to a dimension rather than distributing attention in a more graded way among features *and* dimensions. The development of this skill during the preschool years may in part be supported by the greater ability to use verbal labels as cues for attention to dimensions.

The ID/ED attentional set shifting task

Recently, there has been a call for extending the focus beyond the DCCS when trying to explain the mechanisms and development of attentional set shifting (Buss & Kerr-German, 2019). This is because the DCCS has some features that cannot (or can only partly) be modified: it relies on verbal instructions of explicit rules, thereby posing non-negligible demands on participants' language comprehension skills and the motivation to apply these rules to their behavior. As described above, in tasks for young children, attention to dimension is cued externally through instructions, unlike in adult attention switching tasks such as the Wisconsin Card Sort Task, where participants must deploy attention to a different dimension following feedback from trial-and-error learning (Benitez et al., 2017; Buss & Kerr-German, 2019; Hanania & Smith, 2010). It should be noted that difficulties with understanding the instructions cannot fully explain children's difficulty in the standard version of the DCCS because even 3-year-olds pass task versions using the total change paradigm or using separated dimensions that comprise the same kind and amount of instructions (Diamond, Carlson, & Beck, 2005; Zelazo et al., 2003). Nevertheless, being able to make maximum use of the verbal instructions to direct attention has been suggested to improve with age; young children find the standard version of the DCCS (no change, merged dimensions) combined with verbal instructions to be difficult, whereas older children succeed. What studies using the DCCS cannot do is isolate the challenge of attentional shifting from the opportunity to make use of verbal cues to overcome this challenge.

Performance in the post-switch phase of the DCCS is classified as either pass or fail. Although this is the result of the fact that scores on the task are usually bimodally distributed (Zelazo, 2006) rather than a deliberate dichotomization, it might make the DCCS less suited to studying individual differences in attentional set shifting. A less polarizing measure including more continuous assessment of performance could better capture variation within age groups and be better suited for individual differences studies (Völter, Tinklenberg, Call, & Seed, 2018). Lastly, the DCCS is not easily translatable to different cultures. There is evidence that 5-year-olds' ability to shift is affected by their access to preschool educational resources and their experience with rule-based games and arbitrary symbol mappings (Legare, Dale, Kim, & Deák, 2018). We propose that the ID/ED task is a promising candidate to provide the features identified above as lacking in existing tasks.

The ID/ED task is based on a two-choice compound discrimination learning task (Brown & Tait, 2015a; Sirois & Shultz, 1998). In the pre-switch phase, participants are presented with pairs of compound stimuli that differ along two dimensions (in our study *cup* and *filling material*; Fig. 1). The goal of the task is to identify which stimulus in the set is associated with a reward (compound discrimination). Participants are administered several trials presenting variations of the same four features; the feature predictive of the reward stays the same throughout the trials (red paper in Fig. 1). In each trial, participants can choose only once (forced choice). If the learning criterion is met, participants move on to the post-switch phase, in which the stimuli can either be exchanged for new exemplars (total change) or stay the same (no change). There are two types of shift that can be administered in the post-switch phase. In the total change paradigm—which is the methodology used in the current study—the dimension that is relevant for predicting the reward in the pre-switch phase (e.g., filling material) can either stay the same for the post-switch phase (intradimensional shift [IDS]) or become the irrelevant dimension, with the dimension that was irrelevant in the pre-switch phase becoming the relevant one (extradimensional shift [EDS]). In the no-change paradigm, these two types of shifts have been referred to as reversal and non-reversal shifts, respectively (for a review, see Sirois & Shultz,

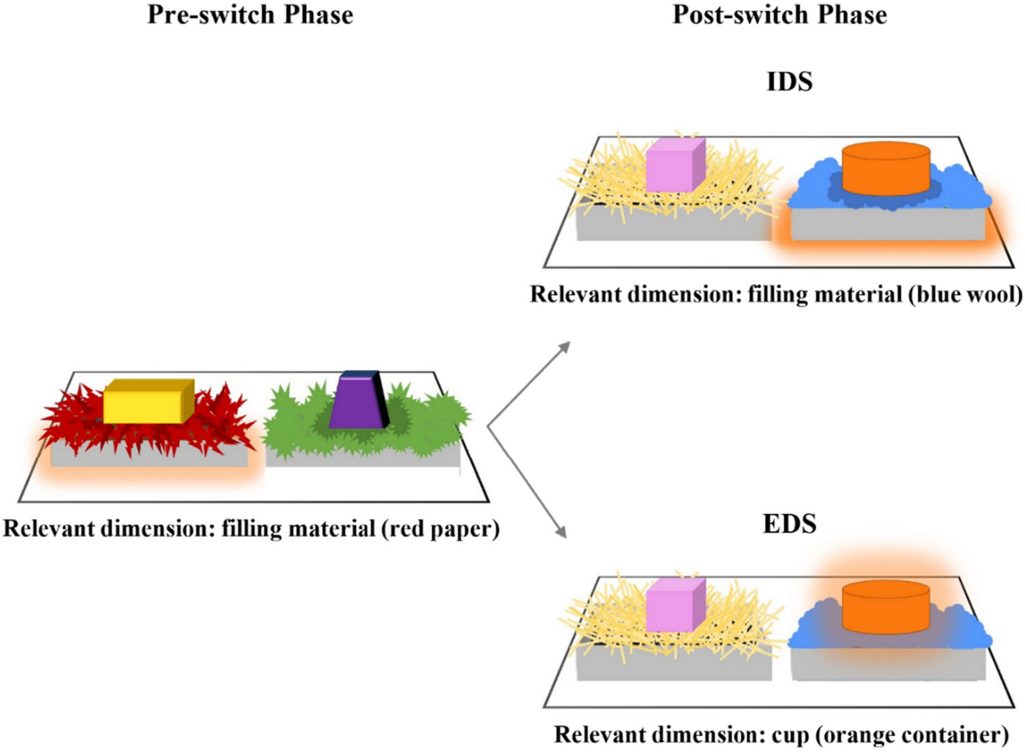

Fig. 1. Example of an Intradimensional/Extradimensional (ID/ED) shift task in a total change version. The two dimensions in the task are filling material and cup. In the pre-switch phase, the relevant dimension is filling material and the rewarded stimulus is the red paper (indicated by the orange glow around the tray with the red paper). The combination of filling material and cup varies between trials, as does the left/right location of the reward. If the learning criterion is met, participants move on to the post-switch phase for which all stimuli are exchanged. In the intradimensional shift (IDS) condition, the dimension that was relevant in the pre-switch phase (here filling material) stays the same (here blue wool relevant). In the extradimensional shift (EDS) condition, the relevant dimension changes (here orange cup). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.).

1998). The phase changes are usually unannounced; that is, in contrast to the DCCS, participants cannot rely on verbal cues but can learn through trial and error (however, note that nonverbal cues of a phase change are given when using total change or partial change designs, when (parts of the) materials are exchanged). The key prediction in the ID/ED paradigm is that if during the pre-switch phase an attentional set is formed toward the relevant dimension, participants in a total change ID/ED task will reach the learning criterion in a subsequent IDS in fewer trials than in an EDS. This pattern of results has been found in human adults and children, rodents, and some monkey species (e.g., macaques) (Brown & Tait, 2015a; Buss, 1953; Eckstrand & Wickens, 1954; Eimas, 1966; Lawrence, 1949; Settlage, Butler, & Odoi, 1956; Slamecka, 1968; Tait, Bowman, Neuwirth, & Brown, 2018; Wolff, 1967). Similarly, for a no-change version, it is predicted that participants would accomplish a reversal (i.e., same dimension relevant in both phases) in fewer trials than a non-reversal. This has been found for human adults (see review by Wolff, 1967) and older children (~10 years; Tighe, Glick, & Cole, 1971), but evidence for younger children, especially preschoolers, has remained inconclusive (cf. Hanania & Smith, 2010; Sirois & Shultz, 1998), partly due to methodological differences between studies (Esposito, 1975; Sirois & Shultz, 1998) and conceptual challenges of the no-change paradigm (see below).

From the 1950s to the 1970s, the ID/ED task was used intensely in various versions to study discrimination learning and attention shifting in humans, but it has not been featured heavily in more recent literature, with its most prevalent use currently being within the Cambridge Neuropsychological Test Automated Battery (CANTAB; Buss, 1953; Caron, 1969; Dickerson, 1966; Dickerson, Wagner, & Campione, 1970; Eimas, 1966; Esposito, 1975; Hanania & Smith, 2010; Kendler & Kendler, 1962; Luciana & Nelson, 2002; Mumbauer & Odom, 1967; Shanab & Yasin, 1979; Shepp & Gray, 1971; Sirois & Shultz, 1998; Slamecka, 1968; Sugimura, 1970; Tighe & Tighe, 2014; Tighe et al., 1971; Völter & Call, 2017; Wolff, 1967). Developmental work using the ID/ED task was pioneered by Kendler and Kendler (1959), who aimed to investigate whether the finding that adults find non-reversals (i.e., a dimension change) to be more difficult than reversals (Buss, 1953) would also hold for young children. Surprisingly, their studies suggested that preschoolers showed the opposite pattern to adults and older children, performing better in non-reversals than in reversals (Kendler, 1979; Kendler, Kendler, & Wells, 1960). Although this was initially explained by young children not yet verbalizing the relevant dimensions, empirical evidence did not support this notion of a requirement of verbal labels (Kendler & Kendler, 1962; Wolff, 1967). Later, Kendler (1979) suggested a tentative theory highlighting the role of the development of selective encoding of information, which could be seen as a precursor of the selective attention theory (see also the Zeaman & House account 1963, 1974, 1984, as cited in Sirois & Shultz, 1998).

However, these early studies have been criticized on three grounds. First, Kendler and Kendler (1959) used a variation of the no-change paradigm that gave children in the non-reversal condition an advantage: in the post-switch phase, the features of the irrelevant dimension did not differ anymore within trials, facilitating a shift in attention towards the new dimension, whereas in the reversal condition no such cue was provided (Eimas, 1965; Sirois & Shultz, 1998). Second, the no-change paradigm as a whole was criticized because using the same stimuli in the pre- and post-switch phases involves continued partial reinforcement of previously relevant features, which complicates comparison between competing theories (Slamecka, 1968). Therefore, subsequent studies used the total change paradigm (see below). Lastly, because some studies could not replicate the finding that non-reversals are easier for young children than reversals (Wolff, 1967), and because the many different versions of shift paradigms make comparisons of these earlier studies difficult (Esposito, 1975; Slamecka, 1968; Wolff, 1967), some have suggested—also assuming a potential "file drawer problem"—that there is no firm evidence yet that young children solve non-reversal shifts more readily than reversals (Sirois & Shultz, 1998).

Studies that have used a no-change paradigm tested children aged 3 to 12 years and showed that children, like adults, performed better in an IDS condition than in an EDS condition of a total change task, providing evidence consistent with the selective attention theory that by 3 years of age at least some children can attend to dimensions (see Table S1 in the online supplementary material). One aspect that has remained unexplored is the developmental trend in performance in ID/ED no-change tasks over the preschool years because in these studies performance was averaged

(Table S1). Therefore, one goal of the current study was to investigate potential age differences in performance on the ID/ED test within the preschool years.

The current study

We created a new ID/ED task (Shifting Tray task) that would be suitable for children aged 3 to 5 years. We used a total change version in order to avoid any confounding effects of partial reinforcement in the post-switch phase (Slamecka, 1968) and to avoid potential floor effects in 3-year-olds. Whereas some ID/ED versions use images of abstract shapes as stimulus materials (see, e.g., the CANTAB), we chose dimensions that children could experience both visually and haptically (Dimension 1: cups of different sizes, shapes, and colors; Dimension 2: substrates/filling materials of different colors and textures), similar to the ID/ED tasks for rodents that use the dimensions *odor* and *digging material* (Brown & Tait, 2015a). Part of the motivation for this design was to make the task easier for preschoolers to learn than the CANTAB, and part of it was to facilitate our comparative studies with nonhuman primates (not described here). Because filling material has not been used before in child studies and might be unusual for children compared with their experience with cups, we included an examination of the effect of relevant dimension on children's learning and switching performance in our analyses (and found no effect).

The dimensions were separated into relatively discrete identities (e.g., a yellow cup on a red substrate) instead of using more integral stimuli as used in the standard version of the DCCS (e.g., a blue star). Versions of the DCCS in which dimensions are separated (i.e., shape depicted on half of the card and color depicted on the other half) aid performance because demands on attentional control are lessened (Doebel & Zelazo, 2015; Landry et al., 2017). We used a minimal amount of instructions to reduce non-shifting-related task demands.

Children started with a pre-switch compound discrimination phase, in which they were repeatedly presented with two trays. Each tray was filled with a different *substrate* (Dimension 1), and on top of each substrate there was a *cup* placed upside down (Dimension 2). The combination of substrates and cups varied between trials. In each trial, children could select one tray with the goal of finding a reward. If children met the learning criterion, they proceeded to either an IDS or an EDS post-switch phase (between-participants design; see Supplementary Video 1 for an example of an EDS condition). We were interested in the proportion of children in each age group reaching the criterion in each phase, whether our task could replicate previous findings that the IDS was easier than the EDS, and whether there were any developmental trends in performance.

Our hypotheses were preregistered and can be found in the preregistration document on the Open Science Framework (OSF) website (https://osf.io/esbqg/?view_only=c154ecc4d-c834d3cbcd954a0ac8f4a69). A copy of them can be found in table format in Table S3. Our main hypotheses concerned the post-switch phase (for the pre-switch hypotheses, see Table S3 or the preregistration). We predicted a nonlinear interaction between age and switch type (IDS/EDS): children would perform better in the IDS condition than in the EDS condition, and this difference in performance would be largest in 4-year-olds compared to 3- and 5-year-olds (Hypothesis 4 in Table S3). Specifically, we based our predictions on the selective attention theory and associated previous findings (Benitez et al., 2017; Hanania, 2010; Hanania & Smith, 2010), which showed that at least some 3-year-olds already attend to the level of dimensions (instead of features) and that with increasing age children focus on dimensions in a more selective manner (all-or-none fashion). In the context of the ID/ED task, this would make them more able to redeploy their attention in the post-switch phase of the EDS condition. We hypothesized that in all age groups children in the IDS condition would on average show better performance than children in the EDS condition, indicating that they selectively attended toward a dimension during the pre-switch phase. This effect would be most pronounced in 4-year-olds compared to 3- and 5-year-olds. The 3-year-olds would be more weakly fixated toward the relevant dimension in the pre-switch phase and thus would show a less pronounced difference between the IDS and EDS conditions. The 5-year-olds should have better selective attention and attentional flexibility skills than 4-year-olds, allowing them to redeploy their attention more quickly, which would also narrow the difference between the IDS and EDS conditions. Therefore, we predicted a nonlinear interaction effect of age and switch type (IDS/EDS) on children's performance, with the difference between conditions being most pronounced in 4-year-olds. It should be noted that the predictions from the other theories we outlined would be very similar: the dynamic neural field theory predicts that older children would perform better than younger children in the post-switch phase due to their stronger neural connections within the frontal cortex and between the frontal and posterior areas. It would also predict that an EDS would be more difficult than an IDS in a total change version of the ID/ED task (A. T. Buss, personal communication October 2020). The attentional inertia theory would suggest that children of all ages would be able to direct their attention toward the level of dimension in the pre-switch phase. In the post-switch phase, older children would be better able to switch than younger children due to their greater inhibitory skills, and performance in the IDS condition would be better than in the EDS condition due to the lower inhibitory demands of an IDS. Thus, both these theoretical frameworks would also lead to the prediction of an effect of condition and age, and an interaction between age and condition could also be made under both theories. Our goal in this study was not to differentiate between these theoretical frameworks but rather to establish a paradigm in which the relative contributions of selective attention and verbal cues could be disambiguated in future work.

Method

Participants

The final sample size was 95 children (49 girls and 46 boys) aged 2 years 11 months to 6 years 0 months (*M* = 54.75 months, *SD* = 10.72) tested between November 2019 and March 2020. There were 1 2-year-old, 26 3-year-olds, 36 4-year-olds, 31 5-year-olds, and 1 6-year-old. In total, 92 children were identified by their parents as White, European, or Caucasian, whereas 1 child was identified as Chinese, 1 as Black, and 1 as multiracial. Target sample size had been 108 children, with 36 each in the age groups of 3, 4, and 5 years. Data collection was planned to cease upon reaching 108 valid participant data points, or at the beginning of March 2020, due to time restraints on the project. The total number of children tested was 114. From these, we needed to exclude 19 children (4 due to experimenter error and 15 due to children losing interest). Of those who dropped out, 6 were 3 years old, 7 were 4 years old, and 2 were 5 years old.

Participants were tested in a science museum in a medium-sized city and at a 1-day event in a small town in Scotland, UK. At both locations, a poster advertisement was placed signaling a participation opportunity for children in the age range. We obtained written informed consent from participants' parents and oral assent from the children prior to the study. Age was the only exclusion criterion. Upon recruitment, participants were randomly allocated to one of eight subconditions within their age group (see below). The study was preregistered in January 2020 (during data collection but before any data analysis was conducted) with the OSF (https://osf.io/esbqg/?view_only=c154ecc4dc834d3cbcd954a0ac8f4a69). Ethical approval was granted by the University of St Andrews, UK, School of Psychology and Neuroscience Ethical Review Committee.

Materials

We used two sets of materials (A and B), each consisting of two transparent square plastic trays (20.3 × 20.3 × 5.1 cm), filled with a substrate and a cup placed upside down on top of the substrate (Fig. 2). In Material Set A, the substrates were red shredded paper and green sand, and the cups were yellow and purple. In Material Set B, the substrates were yellow Easter basket filling material and blue cotton wool, and the cups were pink and orange. A cardboard box was used as an occluder to hide the re-baiting process. All testing sessions were video-recorded. Pens and coloring books were available and set nearby for parents with other children to sit while their children participated.

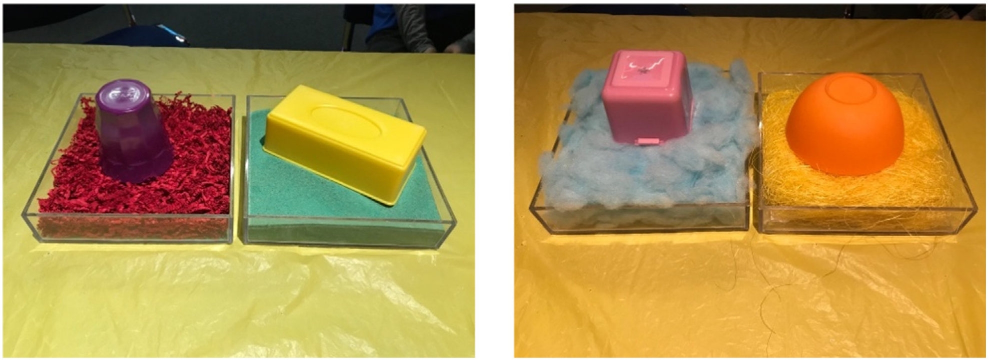

Fig. 2. Materials used in the Shifting Tray task. Left: Material Set A. Right: Material Set B.

Design and procedure

All children were tested by the same female experimenter. Upon recruitment, children were randomly allocated to one of two conditions (IDS or EDS). Within each condition the relevant start dimension could be substrate or cup, and within each start dimension group the material set for the pre-switch phase was either A or B. Material set conditions were added to counterbalance for order effects across the two sets of materials. Allocation to these subgroups was counterbalanced (Table S2).

For both the pre- and post-switch phases, a maximum of 36 trials could be administered. For ease of administration, the 36 trials were split into three blocks of 12 trials. In each block, the four possible stimulus arrangements (see Fig. S1 in the supplementary material) occurred equally often (i.e., three times), with no cup or material appearing on the same side for more than 3 consecutive trials. Within each subgroup, children were allocated to one of five different trial orders for counterbalancing purposes. Gender was also counterbalanced across conditions.

After being briefed and having had questions answered, children and parents were asked whether they would like to participate and, if so, parents were asked to fill in a consent form. Children were told that they could stop at any time. Children then were shown the selection of stickers they might win (e.g., animals, vehicles, stars) and were asked to choose which they would like to collect. They were given an envelope to either stick their stickers on or place them within to take home. The experimenter then explained the game to the children: "In this game you can win lots of stickers. I am going to show you my game, and you can decide whether you want to play. You can stop at any time; you can just let me know. We are going to play several rounds. In each round, I am going to hide a sticker somewhere in there [gestures to trays]. Your task is to try and find the sticker, and each time you can just look once. Would you like to have a go?".

The experimenter placed the occluder to conceal the trays. The trays and cups were organized and a sticker was hidden according to the predetermined order on the coding sheet. Then the experimenter removed the occluder, and children could choose one tray. If children chose successfully, they were congratulated and they took the sticker. If they were unsuccessful, the experimenter said "Oh no! Let's try again," "Never mind," "Oh well, we have lots more chances," or some variations of this. Then the next trial started.

The learning criterion was met once children scored 6 consecutive correct trials in Block 1 (the probability of reaching this by chance is < .05 [0.5^6 =.016]) or once they scored any 9 of 12 trials (75%) correct in Blocks 2 or 3. The pre-switch phase ceased at the end of the block in which the learning criterion was met, if children refused to continue, or if 36 trials were administered without children having met the learning criterion.

If children met the pre-switch learning criterion, they moved on to the post-switch phase. Depending on the condition to which children had been allocated, the relevant dimension in the post-switch phase was either the same as in the pre-switch phase (IDS) or the previously irrelevant one (EDS). The occluder was placed on the table, and children were asked to count how many stickers they had already won. In the meantime, the second material set was brought to the table and the sticker was hidden according to the predetermined order. The post-switch phase was administered in the same way as the pre-switch phase, with the same learning criterion procedures. Testing ceased when children met the criterion, reached the maximum number of 36 trials, or refused to continue.

At the end, children were asked some questions to explore their ability to articulate the rules of the game as an indication of whether they had used inner speech or abstractly understood the game. More information on this procedure and the results can be found in the supplementary material.

Scoring and analysis

Children's responses were live-coded. For each trial, it was coded whether children obtained a sticker (success: yes/no). For pre- and post-switch (if applicable) phases separately, we calculated the total number of administered trials, the number and percentage of successful trials, and the number of trials needed to reach criterion. Data from all but 2 children were coded by a second coder, blind to the hypotheses, to establish inter-rater reliability (for 1 child the video started too late so that coding offline was impossible, and for another child there was no video recording). We found mismatches in the coding for 8 children. Inter-observer reliability was calculated for the variables success, number of administered trials, number of errors, and number of trials after which criterion was reached and was found to be very high (Table S4). All mismatches were double-checked and corrected (6 were mistakes in the original coding and 2 were mistakes in the reliability coding), resulting in perfect agreement.

We investigated the effect of age, relevant dimension, and condition (for post-switch phase only) on children's performance. Our hypotheses focused on the effect of these variables on the number of trials needed to reach criterion. However, when analyzing the data, we decided to also conduct trial-by-trial analyses with success as dependent variable (DV) because those included the entire sample and not just those children who reached criterion. Below we present the models using success as the DV; the models with the number of trials needed to reach criterion are presented in the supplementary material and yield the same qualitative conclusions about our hypotheses.

We investigated whether children's performance in the pre-switch phase was affected by age and relevant dimension. We fitted a generalized linear mixed model (GLMM; Baayen, 2008) with binomial error structure and logit link function on our trial-by-trial data using R Version 3.6.1 (R Core Team, 2020) and the function *glmer* of the R package *lme4* Version 1.1–21 (Bates, Mächler, Bolker, & Walker, 2015). We included success (yes/no) as a DV, age in months (z-transformed to a mean of 0 and a standard deviation of 1 for easier interpretation of the estimates), relevant dimension (cup/substrate), and the interaction between age and relevant dimension as test predictors and included trial number as the control predictor. These predictors were entered as fixed effects. We included a random intercept for participant ID, and to keep the Type I error rate at the nominal level of 5% (Barr, Levy, Scheepers, & Tily, 2013; Schielzeth & Forstmeier, 2009) we included a random slope for trial number. Initially, we included the correlation between the random intercept and slope. However, because this resulted in a singular fit—that is, because the correlation between participant ID and trial number turned out to be unidentifiable as indicated by a correlation parameter of 1 (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017)—we removed the correlation from the model. The model consisted of 2616 observations from 95 children.

Model stability was assessed by comparing the estimates obtained from the model based on all data with those obtained from models with the levels of the random effects excluded one at a time (Nieuwenhuis, te Grotenhuis, & Pelzer, 2012). To rule out collinearity, we determined the variance inflation factors (VIFs; Field, 2005) using a standard linear model excluding the random effects. The model proved to be fairly stable (Table S5; but note that when assessing model stability, we found that several of the models did not converge, likely due to the small sample size). There were no issues of collinearity with this model (VIFs: age = 1.01, relevant dimension = 1.01, trial number = 1.00).

For the post-switch phase, to analyze the potential effects of age, relevant dimension, and condition, we fitted another GLMM with success (yes/no) as the DV, age in months (z-transformed), the polynomial of age in months (z-transformed), relevant dimension (cup/substrate), condition (IDS/

EDS), and the three-way interaction of relevant dimension, condition, and the polynomial of age as test predictors and trial number as the control predictor. These predictors were entered as fixed effects. We included a random intercept for participant ID and a random slope for trial number. However, because this resulted in a singular fit, we first excluded the correlation between the random effects and then (because this resulted in a nonconverging model) removed the random slope from the model. The resulting model consisted of 996 observations from 50 children. There were no issues with collinearity (VIFs: age = 1.05, age^2 = 1.05, relevant dimension = 1.10, condition = 1.04, trial number = 1.09) or with model stability (Fig. S2).

Effect sizes for the entirety of the fixed effects and the entirety of the fixed and random effects were obtained using the function *r.squaredGLMM* of the package *MuMIn* (Bartoń, 2020). As an overall test of the effect of the test predictors, we compared each full model with a null model lacking the test predictors but keeping trial number and the same random effects structure as the full model (Forstmeier & Schielzeth, 2011) using a likelihood ratio test (Dobson, 2002). The *p* values for the individual effects were based on likelihood ratio tests comparing the full model with the respective reduced models using the R function *anova* with argument *test* set to *Chisq* (Barr et al., 2013). All data and the analysis script can be found on the Open Science Framework website https://osf.io/pcs8b/?view_only=f1b1296aca5d41d3a17cf057a5f6cb4c.

Results

Pre-switch phase

There were 46 children (20 boys and 26 girls) in the EDS condition and 49 children (26 boys and 23 girls) in the IDS condition. The distribution of boys and girls did not differ between conditions, chi-square test, $\chi^2(1) = 0.531$, $p = .466$. As expected, the number of children who reached the pre-switch learning criterion in the IDS (n = 26; 53%) did not differ from the number of children who reached the criterion in the EDS condition (n = 24; 52%); neither did the proportion of correct trials differ between the IDS (M = .63, SD = .17) and EDS condition (M = .64, SD = .15), two-samples indepen-

Table 1
Summary of key findings for the pre-switch phase split by age group.

	Age group		
	3 years (*n* = 27)	4 years (*n* = 36)	5 years (*n* = 32)
Mean proportion of correct trials (± *SD*)	.62 ± .19 range = .22–1.00	.65 ± .15 range = .42–.92	.63 ± .15 range = .36–.92
Mean proportion of correct trials if relevant dimension cup (± *SD*)	.60 ± .20 range = .22–1.00	.67 ± .14 range = .42–.92	.60 ± .14 range = .36–.92
Mean proportion of correct trials if relevant dimension substrate (± *SD*)	.64 ± .18 range = .30–1.00	.63 ± .15 range = .42–.92	.67 ± .16 range = .44–.92
Number of administered trials (*M* ± *SD*)	12 trials: 6 24 trials: 5 36 trials: 16 28.44 ± 10.07	12 trials: 12 24 trials: 6 36 trials: 18 26.00 ± 10.92	12 trials: 8 24 trials: 4 36 trials: 20 28.50 ± 10.45
Criterion met	13 (48%)	20 (55%)	17 (53%)
Mean number of trials after which criterion reached (± *SD*)	18.31 ± 10.20 range = 6–36	16.05 ± 9.35 range = 6–36	20.23 ± 11.33 range = 7–36
Mean proportion of correct trials if criterion met (± *SD*)	.75 ± .16 range = .53–1.00	.75 ± .11 range = .54–.92	.74 ± .11 range = .58–.92

Note. The 1 2-year-old was added to the category of 3-year-olds, and the 1 6-year-old was added to the category of 5-year-olds.

Table 2
Summary of key findings for the post-switch phase split by age group.

	Age group					
	3 years (*n* = 13)		4 years (*n* = 20)		5 years (*n* = 17)	
	IDS (*n* = 7)	EDS (*n* = 6)	IDS (*n* = 9)	EDS (*n* = 11)	IDS (*n* = 10)	EDS (*n* = 7)
Mean proportion of correct trials (± *SD*)	.85 ± .08 range = .75–1.00	.72 ± .17 range = .54–.92	.85 ± .14 range = .53–1.00	.67 ± .17 range = .33–.92	.79 ± .17 range = .50–1.00	.70 ± .13 range = .46–.83
Mean proportion of correct trials if relevant dimension cup (± *SD*)	.87 ± .11 range = .79–1.00	.80 ± .21 range = .55–.92	.92 ± .07 range = .83–1.00	.52 ± .13 range = .33–.69	.72 ± .17 range = .50–.92	.66 ± .15 range = .46–.80
Mean proportion of correct trials if relevant dimension substrate (± *SD*)	.83 ±.07 range =.75–.92	.64 ± .10 range = .54–.75	.81 ± .17 range = .53–.92	.80 ± .08 range = .71–.92	.87 ± .16 range = .58–1.00	.75 ± .08 range = .67–.83
Number of administered trials (*M* ± *SD*)	12 trials: 5 24 trials: 2 36 trials: 0 15.43 ± 5.85	12 trials: 2 24 trials: 3 36 trials: 1 22.00 ± 9.03	12 trials: 7 24 trials: 1 36 trials: 1 16.00 ± 8.48	12 trials: 5 24 trials: 2 36 trials: 4 22.91 ± 11.33	12 trials: 6 24 trials: 2 36 trials: 2 19.20 ± 10.12	12 trials: 1 24 trials: 5 36 trials: 1 24.00 ± 6.93
Criterion met	7 (100%)	5 (83%)	8 (89%)	8 (73%)	8 (80%)	7 (100%)
Mean number of trials after which criterion reached (± *SD*)	12.28 ± 6.85 range = 6–22	16.60 ± 8.38 range = 7–24	8.87 ± 4.94 range = 6–21	16.50 ± 9.93 range = 8–35	11.00 ± 6.84 range = 6–22	22.14 ± 7.38 range = 8–33
Mean proportion of correct trials if criterion met (± *SD*)	.85 ± = .08 range = .75–1.00	.75 ± .17 range = .54–.92	.89 ± .07 range =. 75–1.00	.75 ± .11 range = .54–.92	.85 ± .13 range = .58–1.00	.70 ± .13 range = .46–.83

Note. IDS, intradimensional shift; EDS, extradimensional shift.

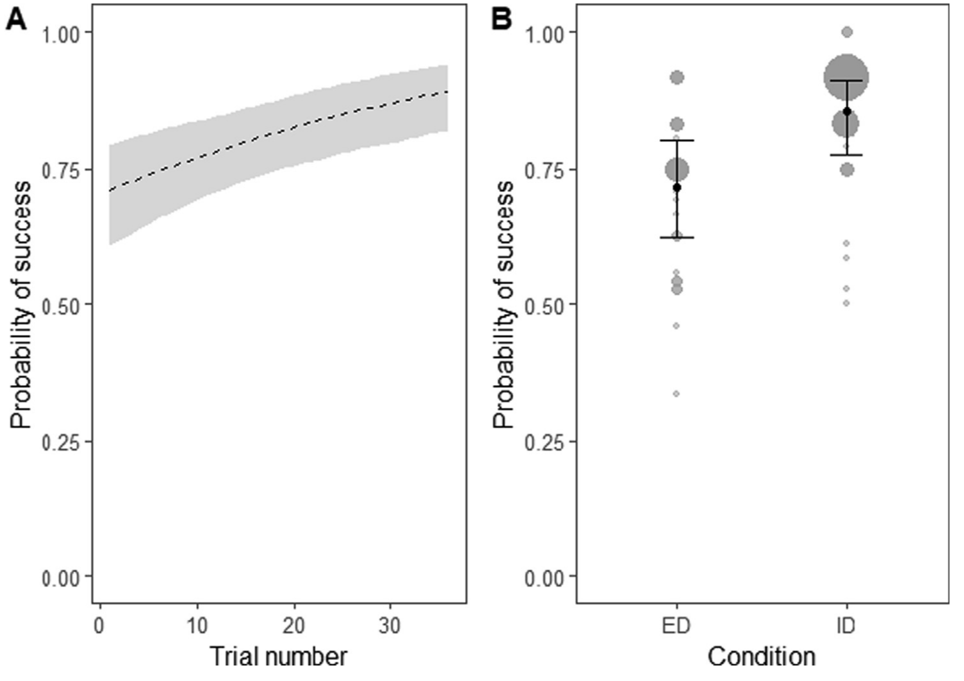

Fig. 3. Effects of trial number (A) and condition (B) on probability of success in the post-switch phase. (A) The dashed line shows the model predictions (with all predictor variables centered), and the gray shaded area shows the bootstrapped 95% confidence interval. (B) shift conditions: ED = extradimensional, ID = intradimensional. The black points show the model predictions (with all predictor variables centered except for condition), the error bars show the bootstrapped 95% confidence interval, and the gray points show the mean performance of each individual. The size of the gray points is proportional to the number of represented individuals.

dent *t* test, $t(92.28) = -0.039$, $p = .969$. All age groups (note that the 1 2-year-old was added to the group of 3-year-olds and the 1 6-year-old was added to the group of 5-year-olds) had a similar proportion of correct trials, ranging from .62 to .65. In all age groups, about half of the children reached the criterion and thus proceeded to the post-switch phase (Table 1). We also found that the proportion of correct trials in conditions where the cup was the relevant dimension was comparable to the proportion of correct trials in conditions where the substrate was the relevant dimension, indicating no preexisting bias to either dimension (see table). We found that together age, relevant dimension, and the interaction between these variables did not explain the data better than a null model containing only the intercept and trial number, $\chi^2(3) = 0.632$, $p = .889$.

Post-switch phase

Of the 95 children, 50 met the pre-switch learning criterion and proceeded to the post-switch stage. In all age groups, the proportion of correct trials was comparably high, with children within each age group performing better in the IDS condition than in the EDS condition (Table 2). Most children reached the learning criterion and did so relatively fast.

The model explained the data significantly better than a null model containing only the intercept and trial number, $\chi^2(11) = 24.357$, $p = .011$. The proportion of variance in the response explained by the entirety of the fixed effects was .10 and that explained by the entirety of the fixed and random effects together was .19. We found that the three-way interaction did not contribute to model fit, $\chi^2(2) = 5.888$, $p = .053$, so we removed it from the model. We then tested the effect of all two-way interactions and found that none of them was significant [Age × Condition: $\chi^2(2) = 0.023$, $p = .988$;

Age × Relevant Dimension: $\chi^2(2)$ = 4.259, p = .119; Condition × Relevant Dimension: $\chi^2(1)$ = 0.802, p = .370], so we removed them from the model. The final model (containing age, the polynomial of age, relevant dimension, condition, trial number, and participant ID) explained the data better than the null model, $\chi^2(4)$ = 13.324, p = .010, and had no model stability issues (Table S6). The proportion of variance in the response explained by the entirety of the fixed effects was .07 and that explained by the entirety of the fixed and random effects together was .20. We found a significant, positive effect of trial number, $\chi^2(1)$ = 13.280, p < .001; with each additional trial, the odds of scoring a trial correct increased by 3%. Above that, there was a significant effect of condition, $\chi^2(1)$ = 9.358, p = .002 (Fig. 3): compared with the EDS condition, the odds of succeeding were 132% larger in the IDS condition. There was no significant nonlinear effect of age, $\chi^2(2)$ = 0.798, p = .671, and no significant effect of relevant dimension, $\chi^2(1)$ = 3.385, p =.066.

Additional exploratory analyses

To get a more detailed picture of children's performance in the post-switch phase, we ran three additional non-preregistered analyses, visualized children's performance by trial blocks and by whether children passed the post-switch learning criterion, investigated whether the difference in the number of correct trials between the pre- and post-switch phases could be predicted by age (result: no), and investigated whether there was a correlation between pre- and post-switch performance split by condition (result: only for IDS). These can be found in the supplementary material.

Discussion

We presented a novel version of the ID/ED task (Shifting Tray task) for studying attentional set shifting in children as young as 3 years. In the pre-switch phase, children were presented with a compound discrimination task in which they needed to attend to a feature within either the dimension *cup* or *substrate*. About half the children in all age groups (3, 4, and 5 years) met the learning criterion and proceeded to the post-switch phase. One reason for the comparatively low level of success in the pre-switch phase of this task compared with the DCCS could be the lack of verbal instructions. The provision of labels increases children's and even infants' selective attention (Althaus & Plunkett, 2015; Perry & Samuelson, 2013) and affects attentional set shifting (Buss & Nikam, 2020; Doebel & Zelazo, 2015). Removing the labels in our task increased the demand on voluntary selective attention for children, which would be in line with previous findings (Buss & Kerr-German, 2019; Buss & Nikam, 2020; Doebel & Zelazo, 2015; Mumbauer & Odom, 1967; Perry & Samuelson, 2013). Another reason could have to do with the specific materials used in this study (e.g., the colors and shapes of the cups). The supplementary material presents an exploratory analysis examining whether some of the children who did not reach the learning criterion in the pre-switch phase might have been "stuck" on a particular nonrewarded stimulus feature, which found that for one third of the children this might have been the case (Table S16). We also found that some children exhibited a side bias (see supplementary material). This is theoretically interesting because it suggests that some children found it difficult to develop an attentional set in the first place. This finding points to an interesting feature inherent to attentional control tasks that do not use explicit rule instructions: In the absence of a rule, learners must set a goal or hypothesis for themselves amid inherently conflicting information (because with each pick of the correct stimulus, the irrelevant dimension is also partially being reinforced). Therefore, the pre-switch phase of such tasks (including the task presented here) resembles somewhat the post-switch phase of other attention shifting tasks, such as the DCCS (Zelazo, 2006) and the SwIFT (an adapted version of the DCCS in which participants need to sort colorful shapes to a single target card per trial; FitzGibbon et al., 2014), in which conflict is being introduced with a rule change. In our task, in the absence of instructions, such a conflict is inherent from the start. Our results relate to the findings by Blakey and colleagues (Blakey & Carroll, 2018; Blakey, Visser, & Carroll, 2016), who found that the post-switch phase of the SwIFT produced a third type of response in addition to switching and perseverating—namely, a mixed response. In the presence of a conflict (but not in the presence of merely distracting information), the youngest children (2.5- to 3-year-olds) struggled with upholding a systematic, rule-governed response. The proportion of mixed responses decreased by 3.5 years of age, when children went through a phase of perseveration, before reaching greater proportions of switching by 4 years of age. It could be possible that in the current study some of the unsuccessful children in the pre-switch phase also showed a mixed response in the face of the inherently conflicting information. The fact that we also found this mixed response in 4- and 5-year-olds could be due to increased task demands because no instructions were provided and children needed to set their own goals and maintain them, which has been shown to be an additional source of difficulty in attention shifting tasks (Cragg & Chevalier, 2012).

If in the future researchers wanted to use the Shifting Tray task and get more children to pass the pre-switch criterion, it might be desirable to increase the salience of the relevant dimension. For example, this could be done by replacing some trials with homogeneous trials in which the irrelevant dimension is held constant (e.g., if material is irrelevant: two different cups sitting on the same material type). A previous pilot study (unpublished) using homogeneous trials indeed resulted in a higher pass rate (84%) in the pre-switch phase (see supplementary material). Another possibility would be to use verbal labels or some other form of contextual cue—yet note that this would probably come with a cost for post-switch performance. However, because we have set up our task as a new test to further examine theoretical questions, we deem the low performance rate in the pre-switch phase to be acceptable.

In the post-switch phase of the current study, children were administered either an EDS or an IDS. Our task was shown to have captured attentional set formation because we could replicate results from previous studies (Table S1) finding that the IDS condition was easier than the EDS condition. This was the case for all age groups, which is also in line with these earlier studies that found the IDS–EDS difference for children aged 3 to 9 years and 12 years and for young adults. These results show that even some of the youngest children's attention was at least partly directed toward dimensions.

With regard to the developmental trajectory, we found no age effect on performance in the post-switch phase. Previous studies using the total change version of the ID/ED task did not test for age effects and often did not present the results by age groups, so this is a new finding within the context of this task. In addition, the studies in Table S1 often differ in stimulus materials, learning criteria, and DVs, making it difficult to identify potential patterns related to age.

One reason for the lack of an age effect in the current study could be that in the design of the task we intended to boost 3-year-olds' performance to avoid floor effects. First, we used a total change version of the paradigm. Total change versions allow a larger proportion of 3- and 4-year-olds to switch because the post-switch phases lack the perceptual conflict that arises in a no-change version. In total change versions of the DCCS, where the stimuli on the cards differ between the pre- and post-switch phases (i.e., same dimensions, different features), pass rates for 3-year-olds ($\sim$73%) are twice as high as in the standard version (36%) (Landry et al., 2017). Second, the use of more separable dimensions in our task (cup and substrate instead of shape and color of a *single* object) possibly also boosted 3-year-olds' performance. Therefore, future studies could examine the impact of these variables, for example, by using partial or no-change versions of this task, which could possibly increase variation in performance across the preschool years.

The lack of labels might be another reason why we did not find any performance differences in the post-switch phase with age. The provision of or instruction to use meaningful labels aids children in attentional set shifting tasks (Buss & Nikam, 2020; Doebel & Zelazo, 2015; Mumbauer & Odom, 1967; Perry & Samuelson, 2013), and the more experience children have with labels for stimuli and dimensions (e.g., the older they are), the better they can use them to focus and shift their attention (Buss & Nikam, 2020). However, despite their arguably greater experience with labels for cups and colors, 5-year-olds were not more able than 3-year-olds to switch their attention between dimensions. They seemingly did not label the different aspects of the task for themselves given that no children used labels for dimensions when prompted to explain how they solved the task. This result is in line with the conclusion drawn from the studies from the 1960s and 1970s that whereas "verbal mechanisms must be considered as an important factor" for performance in ID/ED tasks (as suggested by the work of the Kendlers), "the principal factors operating in the shift process *in general* are probably attentional in nature" (Wolff, 1967, p. 403). Future studies including older children could investigate the onset of the spontaneous use of verbal labels.

In summary, the Shifting Tray task was shown to be a valid test of attentional set shifting in preschoolers (because it detects switching costs by demonstrating an ID/ED difference) and thus extends the still limited set of available set shifting tasks for this age group. With its flexible design, the task provides an alternative to the DCCS because it can be used without an explicit rule structure or labels. Importantly, these could be reintroduced to the task to investigate the role played by this variable. In contrast to the DCCS, we found no developmental increase in switching performance. However, it would be important to examine a wider range of versions of this task (e.g., using labels, using partial or no-change versions) to examine whether this affects the developmental trajectory of performance in the post-switch phase.

We propose that further studies using the ID/ED task (along with other attentional set shifting tasks) would contribute to gaining insights into the mechanisms underlying the development of attentional set shifting and the factors affecting it. Due to their flexible nature, ID/ED tasks can be used to test predictions of current theories and to study the impact of verbal labels on attentional set shifting. Due to the minimal amount of instructions and the use of less abstract stimuli than, for example, the DCCS and the ID/ED task of the CANTAB, the Shifting Tray task also has the potential to be more easily adapted to non-Westernized populations. Most research on EFs (including the current study) has been done on samples from WEIRD (Westernized, highly Educated, Industrialized, Rich, and Democratic) backgrounds (Apicella, Norenzayan, & Henrich, 2020; Henrich, Heine, & Norenzayan, 2010) and thus are limited in their generalizability to non-Westernized populations (Holding et al., 2018; Legare et al., 2018). Finally, the task can also be used to study attentional set shifting comparatively because it can be used with nonhuman animals without major changes (Brown and Tait, 2015a). In fact, the Shifting Tray task has recently been included in a test battery measuring EFs in children and chimpanzees (Völter et al., 2022). Thus, the Shifting Tray task has the potential to be a useful tool to further understand the development and evolution of attentional control.

Acknowledgments

We thank all the children and their parents for participation in the study; Geraldine Brown and the Dundee Science Centre for facilitating data collection; Vvidhi Agrawwal, Marcio Andrade, Mei Angwin, Avanti Chopra, Harry Siderfin, and Wendi Zhang for their help with data collection; Lisa Duncan and Zsuzsa Lugosi for reliability coding; Mio Hisada, Ilkay Makhotina, Lesley Motherwell, Lina Soderstrom, Alex Backler, Junie Wee, Nitika Singh, and Anastasia Seccombe for their help with coding; and three anonymous reviewers for comments on earlier versions of the manuscript. The research of A.M.S. was supported by an "INQMINDS" European Research Council (ERC) Starting Grant (SEP-210159400). The funding source had no involvement in study design, collection, analysis, interpretation of the data, writing the manuscript, or the decision to submit the manuscript for publication.

Author contributions

E.R.: conceptualization, methodology, formal analysis, data curation, writing–original draft, writing–review & editing, revision, visualization, and supervision; C.J.V.: conceptualization, methodology, writing–review & editing, revision, and visualization; J.C-M.: conceptualization, methodology, formal analysis, investigation, data curation, and writing–review & editing; J.C.: conceptualization, writing–review & editing, and revision; A.M.S.: conceptualization, methodology, resources, writing–review & editing, revision, supervision, project administration, and funding acquisition.

Appendix A. Supplementary data

Supplementary information to this article can be found online at https://doi.org/10.1016/j.jecp.2022.105428. All data and the analysis script can be found on the Open Science Framework website https://osf.io/pcs8b/?view_only=f1b1296aca5d41d3a17cf057a5f6cb4c.

References

Althaus, N., & Plunkett, K. (2015). Categorization in infancy: Labeling induces a persisting focus on commonalities. *Developmental Science, 19*(5), 1–11. https://doi.org/10.1111/desc.12358.

Apicella, C., Norenzayan, A., & Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior, 41*, 319–329.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Bardikoff, N., & Sabbagh, M. (2017). The differentiation of executive functioning across development: Insights from developmental cognitive neuroscience. In N. Budwig, E. Turiel, & P. Zelazo (Eds.), *New perspectives on human development* (pp. 27–46). Cambridge, UK: Cambridge University Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Bartoń, K. (2020). *MuMIn: Multi-model inference.* https://cran.r-project.org/web/packages/MuMIn/index.html

Bates, D. M., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01.

Benitez, V. L., Vales, C., Hanania, R., & Smith, L. B. (2017). Sustained selective attention predicts flexible switching in preschoolers. *Journal of Experimental Child Psychology, 156*, 29–42.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81*, 1641–1660.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*, 327–336.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647–663.

Blakey, E., & Carroll, D. J. (2018). Not all distractions are the same: Investigating why preschoolers make distraction errors when switching. *Child Development, 89*, 609–619.

Blakey, E., Visser, I., & Carroll, D. J. (2016). Different executive functions support different kinds of cognitive flexibility: Evidence from 2-, 3-, and 4-year-olds. *Child Development, 87*, 513–526.

Brown, V. J., & Tait, D. S. (2015a). Attentional set-shifting across species. In T. W. Robbins & B. J. Sahakian (Eds.). *Translational neuropsychopharmacology (Current Topics in Behavioral Neurosciences)* (Vol. 28, pp. 363–395). Cham, Switzerland: Springer.

Brown, V. J., & Tait, D. S. (2015b). Behavioral flexibility: Attentional shifting, rule switching, and response reversal. In I. P. Stolerman & L. H. Price (Eds.), *Encyclopedia of psychopharmacology* (pp. 264–269). Berlin: Springer.

Bull, R., Espy, K. A., & Senn, T. E. (2004). A comparison of performance on the Towers of London and Hanoi in young children. *Journal of Child Psychology and Psychiatry, 45*, 743–754.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205–228.

Burgess, P. W. (1997). Theory and methodology in executive function research. In P. Rabbitt (Ed.), *Theory and methodology of frontal and executive function* (pp. 81–116). New York: Psychology Press.

Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science, 29*, 624–630.

Buss, A. H. (1953). Rigidity as a function of reversal and non-reversal shifts in the learning of successive discriminations. *Journal of Experimental Psychology, 45*, 75–81.

Buss, A. T., & Kerr-German, A. (2019). Dimensional attention as a mechanism of executive function: Integrating flexibility, selectivity, and stability. *Cognition, 192* 104003.

Buss, A. T., & Nikam, B. (2020). Not all labels develop equally: The role of labels in guiding attention to dimensions. *Cognitive Development, 53* 100843.

Buss, A. T., & Spencer, J. P. (2014). The emergent executive: A dynamic field theory of the development of executive function. *Monographs of the Society for Research in Child Development, 79*(2, Serial No. 313).

Buss, A. T., & Spencer, J. P. (2018). Changes in frontal and posterior cortical activity underlie the early emergence of executive function. *Developmental Science, 21* e12602.

Caron, A. J. (1969). Discrimination shifts in three-year-olds as a function of dimensional salience. *Developmental Psychology, 1*, 333–339.

Chevalier, N., & Blaye, A. (2008). Cognitive flexibility in preschoolers: The role of representation activation and maintenance. *Developmental Science, 11*, 339–353.

Cragg, L., & Chevalier, N. (2012). The processes underlying flexibility in childhood. *Quarterly Journal of Experimental Psychology, 65*, 209–232.

Deák, G. O. (2000). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development, 1*, 157–191.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168.

Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In P. McCardle, L. Freund, & J. A. Griffin (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research* (pp. 11–43). Washington, DC: American Psychological Association.

Diamond, A., Carlson, S. M., & Beck, D. M. (2005). Preschool children's performance in task switching on the Dimensional Change Card Sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology, 28*, 689–729.

Dickerson, D. J. (1966). Performance of preschool children on three discrimination shifts. *Psychonomic Science, 4*, 417–418.

Dickerson, D. J., Wagner, J. F., & Campione, J. (1970). Discrimination shift performance of kindergarten children as a function of variation of the irrelevant shift dimension. *Developmental Psychology, 3*, 229–235.

Dobson, A. J. (2002). *An introduction to generalized linear models*. Boca Raton, FL: Chapman & Hall/CRC Press.

Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review, 38*, 241–268.

Eckstrand, G. A., & Wickens, D. D. (1954). Transfer of perceptual set. *Journal of Experimental Psychology, 47*, 274–278.

Eimas, P. D. (1965). Comment: Comparisons of reversal and nonreversal shifts. *Psychonomic Science, 3*, 445–446.

Eimas, P. D. (1966). Effects of overtraining and age on intradimensional and extradimensional shifts in children. *Journal of Experimental Child Psychology, 3*, 348–355.

Esposito, N. J. (1975). Review of discrimination shift learning in young children. *Psychological Bulletin, 82*, 432–455.

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology, 13*, 495–499.

Espy, K. A., & Cwik, M. F. (2004). The development of a trail making test in young children: The TRAILS-P. *The Clinical Neuropsychologist, 18*, 411–422.

Field, A. (2005). *Discovering statistics using SPSS*. Thousand Oaks, CA: Sage.

Fisher, A. V. (2019). Selective sustained attention: A developmental foundation for cognition. *Current Opinion in Psychology, 29*, 248–253.

FitzGibbon, L., Cragg, L., & Carroll, D. J. (2014). Primed to be inflexible: The influence of set size on cognitive flexibility during childhood. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00101.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology, 65*, 47–55.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*, 483–527.

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*, 31–60.

Hanania, R. (2010). Two types of perseveration in the Dimension Change Card Sort task. *Journal of Experimental Child Psychology, 107*, 325–336.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science, 13*, 622–635.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61–135.

Henry, L. A., & Bettenay, C. (2010). The assessment of executive functioning in children. *Child and Adolescent Mental Health, 15*, 110–119.

Holding, P., Anum, A., van de Vijver, F. J. R., Vokhiwa, M., Bugase, N., Hossen, T., Makasi, C., Baiden, F., Kimbute, O., Bangre, O., Hasan, R., Nanga, K., Sefenu, R. P. S., A-Hayat, N., Khan, N., Oduro, A., Rashid, R., Samad, R., Singlovic, J., Faiz, A., & Gomes, M. (2018). Can we measure cognitive constructs consistently within and across cultures? Evidence from a test battery in Bangladesh, Ghana, and Tanzania. *Applied Neuropsychology: Child, 7*(1), 1–13. https://doi.org/10.1080/21622965.2016.1206823.

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30*, 190–200.

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*, 1147–1185.

Kendler, H. H., & Kendler, T. S. (1962). Vertical and horizontal processes in problem solving. *Psychological Review, 69*, 1–16.

Kendler, T. S. (1979). The development of discrimination learning: A levels-of-functioning explanation. In H. W. Reese & L. P. Lipsitt (Eds.). *Advances in child development and behavior* (Vol. 13, pp. 83–117). New York: Academic Press.

Kendler, T. S., & Kendler, H. H. (1959). Reversal and nonreversal shifts in kindergarten children. *Journal of Experimental Psychology, 58*, 56–60.

Kendler, T. S., Kendler, H. H., & Wells, D. (1960). Reversal and nonreversal shifts in nursery school children. *Journal of Comparative and Physiological Psychology, 53*, 83–88.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–467.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the Dimensional Change Card Sorting task. *Developmental Science, 8*, 44–56.

Landry, O., Al-Taie, S., & Franklin, A. (2017). 3-year-olds' perseveration on the DCCS explained: A meta-analysis. *Journal of Cognition and Development, 18*, 419–440.

Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology, 39*, 770–784.

Legare, C. H., Dale, M. T., Kim, S. Y., & Deák, G. O. (2018). Cultural variation in cognitive flexibility reveals diversity in the development of executive functions. *Scientific Reports, 8*, 16326.

Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology, 22*, 595–624.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. M. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language, 94*, 305–315.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.

Mumbauer, C. C., & Odom, R. D. (1967). Variables affecting the performance of preschool children in intradimensional, reversal, and extradimensional shifts. *Journal of Experimental Psychology, 75*, 180–187.

Munakata, Y., Morton, J. B., & Yerys, B. E. (2003). Children's perseveration: Attentional inertia and alternative accounts. *Developmental Science, 6*, 471–473.

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *The R Journal, 4*(2), 38–47.

Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the Dimensional Change Card Sorting task? *Infant and Child Development, 11*, 93–105.

Perry, L., & Samuelson, L. (2013). The role of verbal labels in attention to dimensional similarity. In M. Knauff (Ed.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, pp. 3217–3222). Austin, TX: Cognitive Science Society.

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Quarterly Journal of Experimental Psychology, 40*, 321–341.
Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology, 20*, 416–420.
Settlage, P., Butler, R., & Odoi, H. (1956). Perseverative interference in monkeys following bilateral removal of the prefrontal areas. *Journal of General Psychology, 54*, 255–262.
Shanab, M. E., & Yasin, A. B. (1979). Intradimensional and extradimensional shifts by Jordanian students. *Journal of General Psychology, 100*, 199–213.
Shepp, B. E., & Gray, V. A. (1971). Some effects of variable-within and variable-between irrelevant stimuli on dimensional learning and transfer. *Journal of Experimental Psychology, 89*, 32–39.
Sirois, S., & Shultz, T. R. (1998). Neural network modeling of developmental effects in discrimination shifts. *Journal of Experimental Child Psychology, 71*, 235–274.
Slamecka, N. J. (1968). A methodological analysis of shift paradigms in human discrimination learning. *Psychological Bulletin, 69*, 423–438.
Smidts, D. P., Jacobs, R., & Anderson, V. (2004). The object classification task for children: A new measure of concept generation and mental flexibility in early childhood. *Developmental Neuropsychology, 26*, 385–401.
Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology, 24*, 279–298.
St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*, 745–759.
Sugimura, T. (1970). Consecutive intradimensional, extradimensional, and control shifts in kindergarten children. *Japanese Psychological Research, 12*(1), 36–43.
Tait, D. S., Bowman, E. M., Neuwirth, L. S., & Brown, V. J. (2018). Assessment of intradimensional/extradimensional attentional set-shifting in rats. *Neuroscience and Biobehavioral Reviews, 89*, 72–84.
Tighe, L. S., & Tighe, T. J. (2014). Overtraining and discrimination shift behavior in children. *Psychonomic Science, 2*, 365–366.
Tighe, T. J., Glick, J., & Cole, M. (1971). Subproblem analysis of discrimination-shift learning. *Psychonomic Science, 24*, 159–160.
Völter, C. J., & Call, J. (2017). Causal and inferential reasoning in animals. In J. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, & T. Zentall (Eds.), *APA handbook of comparative psychology: Perception, learning, and cognition (APA Handbooks in Psychology)* (pp. 643–671). Washington, DC: American Psychological Association.
Völter, C. J., Tinklenberg, B., Call, J., & Seed, A. M. (2018). Comparative psychometrics: Establishing what differs is central to understanding what evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*, 20170283.
Völter, C. J., Reindl, E., Felsche, E., Civelek, Z., Whalen, A., Lugosi, Z., Duncan, L., Herrmann, E., Call, J., & Seed, A. M. (2022). The structure of executive functions in preschool children and chimpanzees. *Scientific Reports, 12*, 6456.
Wolff, J. L. (1967). Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin, 68*, 369–408.
Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297–302.
Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), 93–119.