

The benefits of immature cognitive control:

How distributed attention guards against learning traps

Running Head: Benefits of Immature Control

Word Count: 8547

Data availability Statement:

Data from this study are available on Open Science Framework at:

https://osf.io/w75he/?view_only=10085d9756d0421e9d450f57c5c6e219

Declarations of interest: none

Research Highlights

- Cognitive control allows one to focus attention selectively. Selective attention has both benefits and costs. One such cost is learned inattention.
- Adults tend to focus attention selectively, whereas young children tend to distribute attention broadly.
- Using eye-tracking analyses, we directly assess how participants' attention allocation patterns during learning affect the tendency to fall into learning traps due to learned inattention.
- Results indicate that young children's tendency to distribute their attention broadly may be adaptive, protecting them from the potential pitfalls of selective attention.

Abstract

Cognitive control allows one to focus their attention efficiently on relevant information while filtering out irrelevant information. This ability provides a means of rapid and effective learning, but using this control also brings risks. Importantly, useful information may be ignored and missed, and learners may fall into “learning traps” (e.g., learned inattention) wherein they fail to realize that what they ignore carries important information. Previous research (e.g., Plebanek & Sloutsky, 2017) has shown that adults may be more prone to such traps than young children, but the mechanisms underlying this difference are unclear. The current study uses eye-tracking to examine the role of attentional control during learning in succumbing to these learning traps. Four-year-old children and adults completed a category learning task in which an unannounced switch occurred wherein the feature dimensions most relevant to correct categorization became irrelevant, and formerly irrelevant dimensions became relevant. After the switch, adults were more likely than children to ignore the new highly relevant dimension and settle on a suboptimal categorization strategy. Further, eye-tracking analyses reveal that greater attentional selectivity during learning (i.e., optimizing attention to focus only on the most relevant sources of information) predicted this tendency to miss important information later. Children’s immature cognitive control, leading to broadly distributed attention, appears to protect them from this trap—although at the cost of less efficient and slower learning. These results demonstrate the double-edged sword of cognitive control and suggest that immature control may serve an adaptive function early in development.

Keywords: cognitive control; adaptive immaturity; selective attention; distributed attention; learning traps; category learning

The benefits of immature cognitive control: how distributed attention guards against learning traps

The world around us contains an almost infinite amount of potentially useful information. However, our limited time and cognitive processing capacity constrain our ability to examine all these sources of information. Therefore, information relevant to our goals must be prioritized, and less relevant information should be filtered out to make efficient use of our time and cognitive resources. One way of prioritizing information is through top-down or endogenous control, allowing for the selective focus of attention on relevant information and away from irrelevant information (Desimone & Duncan, 1995; Soto, Hodsoll, Rotshtein & Humphreys, 2008). The ability to selectively attend enables quick and efficient learning, but it also relies on previous experience and knowledge to assess what *is* and what *is not* relevant.

In contrast to adults, young children may lack both the knowledge of what is relevant and the ability to ignore what is irrelevant. Across childhood, they acquire this necessary experience and knowledge, while their ability to control their attention is also developing (Leclercq & Siéroff, 2013; Shimi, Nobre, Astle & Scerif, 2014; Wainwright & Bryson, 2005). Top-down control of attention develops gradually throughout childhood, possibly even into young adulthood (Goldberg, Maurer & Lewis, 2001; Jakobsen, Frick & Simpson, 2013; Schul, Townsend & Stiles, 2003). Important developmental transitions in endogenous attentional control appear to occur around the ages of six or seven (Munakata, Snyder & Chatham, 2012), but such control continues developing—not reaching adult-like levels until around ages nine to eleven (Goldberg, Maurer & Lewis, 2001; Leclercq & Siéroff, 2013; Pearson & Lane, 1990; Schul, Townsend & Stiles, 2003; Wainwright & Bryson, 2005), or possibly even later (Wong-Kee-You, Tsotsos & Adler, 2019). Consequently, endogenous control of attention is relatively immature in young children (Hanania

& Smith, 2010; Plude, Enns, & Brodeur, 1994), and children under 6 years of age have difficulty prioritizing relevant over irrelevant information. As a result, young children often fail to attend selectively, tending to distribute their attention broadly (e.g., Best, Yim, & Sloutsky, 2013; Deng & Sloutsky, 2015a, 2015b, 2016; Plebanek & Sloutsky, 2017; Smith & Kemler, 1977).

Benefits and Costs of Distributed vs. Selective Attention

It is easy to see that selective and distributed attention have different costs and benefits. For example, when searching for a friend wearing a red jacket in a crowd, selective attention would allow a person to focus exclusively on people wearing red jackets. In contrast, distributed attention would not limit search to red color. As a result, search under selective attention would be vastly more efficient than search under distributed attention. However, this is true only if one's knowledge is accurate. Imagine that the lost friend decided to take the red jacket off and is now wearing a blue sweater. Under these circumstances, a person attending selectively may be less likely to find the friend than a person who distributes their attention broadly.

More generally, cognitive control allows one to flexibly adapt their attention depending on goals and task demands, by either focusing it selectively or distributing it broadly. As mentioned above, both selective and distributed attention have benefits as well as costs.

Attentional selectivity is often beneficial in that it leads to faster, more efficient processing of the attended information, supporting the ability to learn and react rapidly. On the other hand, selective attention also has costs (Blanco & Sloutsky, 2019; Hoffman & Rehder, 2010; Plebanek & Sloutsky, 2017; Rich & Gureckis, 2018). The short-term cost of selective attention is that unattended information is missed, which can be detrimental if this information turns out to be important later. Perhaps more importantly, there are also potential long-term effects of selective attention, sometimes referred to as "learning traps" (Rich & Gureckis, 2018). One such trap is

learned inattention (see Hoffman & Rehder, 2010 for a review), wherein a source of information that is learned to be irrelevant continues to be ignored in the future, leading to difficulty learning from that information source if it later becomes useful (Kruschke & Blair, 2000).

Broad distribution of attention, conversely, leads to less efficient learning and performance on cognitive tasks (see Desimone & Duncan, 1995, for review) due to the increased time spent processing information that is likely less relevant to the decision. Additionally, irrelevant information may distract one from their goals. The positive side of this tradeoff is that some of that information may be useful later if the conditions or goals change.

Recent research has highlighted how these benefits and costs manifest across development due to differences in attention allocation between young children and adults. For example, Plebanek and Sloutsky (2017) presented 4- to 5-year-old children and adults with a change detection task, similar to the task used by Rock and Gutman (1981), in which participants observed two overlaid shapes and their attention was cued to one of the shapes. The shapes then disappeared, and after a delay, test shapes appeared on the screen. These shapes could be the same as the studied shapes, or they could differ on either the cued or uncued shape. Participants were instructed to respond “same” if the new shapes were identical to the studied shapes, or “different” otherwise. The results revealed an interesting dissociation wherein adults were somewhat more accurate at detecting changes in the cued shape (presumably due to selectively attending to it), but children were substantially better at detecting changes in the uncued shape (presumably due to distributing their attention to both shapes). The study also demonstrated analogous results in visual search, wherein children outperformed adults during a surprise subsequent memory test for features that were irrelevant to the main search task (Plebanek & Sloutsky, 2017).

Across several studies, Deng and Sloutsky (2015a, 2016) found similar effects during category learning. In their studies, the category structure that participants learned was designed to have one perfectly diagnostic feature dimension that predicted the correct category label and several probabilistically predictive feature dimensions (that usually, but not always, predicted the correct category label). Adults quickly discovered the diagnostic dimension, used it almost exclusively when making category judgments, and remembered its values much better than those of the other dimensions. Four- to five-year-old children, on the other hand, relied on multiple dimensions to classify category members and had relatively good memory for all the features, exhibiting better memory for probabilistic features than adults.

In another category learning study (that has design similar to that of the current study), Blanco and Sloutsky (2019) examined learned inattention in preschool-aged children and adults. In addition, to the deterministic and probabilistic feature dimensions, the stimuli used in this research had also an irrelevant dimension that was not predictive of the category label. After initial learning, a diagnostic dimension of the stimuli and an irrelevant dimension swapped roles. Adults were less likely than children to discover the value of the newly diagnostic dimension and to utilize it for category judgments in the second half of the experiment. Instead, it seems that adults continued ignoring the dimension after learning it was irrelevant in the first half of the experiment, suggesting that they experienced more learned inattention than young children.

The studies described above suggest that children sometimes notice things that adults miss, which can benefit them, and it is hypothesized that this benefit is a result of developmental differences in attentional focus and ultimately in the control of attention. However, in these studies, the role of attentional control was inferred from behavioral patterns rather than measured directly. In the current study, we attempt to examine a more direct link between attention allocation and

patterns of behavior by directly measuring (via eye-tracking) attention allocation during category learning and using these data to predict later categorization responses.

The Current Study

The previous studies described above relied on group level differences in performance and indirect measures of attention that suggested differences in attention allocation between the age groups, but attention was inferred from performance rather than assessed directly. The current study adapted the task design used by Blanco and Sloutsky (2019) while incorporating eye-tracking to directly evaluate the connection between attentional allocation and learned inattention. Although visual fixations are not equivalent to attention (as participants can attend covertly), they have been used as a useful and informative proxy of attention.

In the current task, four- to five-year-old children and adults learned categories, while their eye gaze was recorded. Children of this age were included in the study because 4-year-olds exhibited categorization behavior that was different from adults and indicative of distributed attention in previous studies (Blanco & Sloutsky 2019; Deng & Sloutsky, 2015a, 2016; Plebanek & Sloutsky, 2017), while older children (6-year-olds) behaved similarly to adults, exhibiting behavior consistent with selective attention (Deng & Sloutsky, 2016).

Participants' goal in the task was to learn to classify novel artificial creatures into two categories. These creatures were composed of different dimensions (e.g., feet, hands, or tail) that take on different feature values which vary in how well they predict the category label (see Figure 1). One dimension was deterministic, as it always perfectly predicted the correct category label. One dimension was completely irrelevant, as it never predicted the category label above the chance level. The rest of the dimensions were probabilistic, as they predicted category labels with some (greater than chance) probability. For example, if the creature's tail was a probabilistic dimension,

a particular type of tail would usually be found in members of category A but occasionally in members of category B (i.e., four out of five times). Because of this category structure, different attention allocation strategies could be used to successfully learn the categories, allowing for variation in the deployment of cognitive control. One effective strategy is to focus selectively on only the deterministic dimension, while another is to pay attention to all dimensions and learn the category structure based on the overall similarity among the category members. Previous studies have indicated that adults typically use the first strategy when learning this type of category structure, whereas young children tend to use the second strategy (Deng & Sloutsky, 2015a, 2016).

To examine the consequences of these different ways of controlling attention, the task included an unannounced switch halfway through the experiment in which the previously perfectly predictive dimension became irrelevant to making accurate categorization judgments, while the dimension that was previously irrelevant became perfectly predictive. If participants optimized attention and learned to ignore the initially irrelevant dimension, they would have difficulty noticing that this dimension has become a valuable source of information, falling into a learning trap due to learned inattention. By contrast, if participants distributed attention, they might avoid learning traps. In addition, instead of assuming that young children (as a group) distribute attention and adults (as a group) optimize attention, we measure individuals' attention allocation directly.

To preview, our results suggest that children more broadly distributed their attention and were less prone to learned inattention compared adults. Furthermore, the level of selectivity during initial learning accurately predicted whether individual participants would succumb to learned inattention.

Method

Participants

The participants were 30 4-5-year-old children (17 females, 13 males, $Mean_{age} = 54$ months, $SD_{age} = 3.66$ months, $Range_{age} = 48-61$ months) and 38 adults (25 females, 13 males, $Mean_{age} = 19$ years, $SD_{age} = 1.87$ years, $Range_{age} = 18-29$ years). An additional 6 children and 3 adults were recruited but excluded due to technical issues (2 children; 2 adults) or failure to complete the task (4 children; 1 adult). Children were recruited from preschools and childcare centers located in middle class neighborhoods near [removed for blind review]. Adults were undergraduate students at [removed for blind review], who participated in exchange for partial course credit. We obtained informed consent for all participants in our study. The study was approved by the Institutional Review Board at [removed for blind review]. Sample sizes were based on those reported in previous publications that use similar tasks (Blanco & Sloutsky, 2019; Deng & Sloutsky 2015a, 2016).

Stimuli

The stimuli were colorful images of creatures composed of seven discrete-valued dimensions (see Figure 1) similar to those used in previous studies (Blanco & Sloutsky, 2019; Deng & Sloutsky, 2015a, 2016). The creatures were divided into two categories, which were referred to in the experiment as “Flurps” and “Jalets”. The stimuli had seven dimensions: antenna, head, body, neck button, hands, feet, and tail. Importantly, one dimension deterministically predicted category membership (i.e., the Deterministic dimension), five dimensions were probabilistically predictive with 80% cue validity (i.e., the Probabilistic dimensions), and one dimension was non-predictive — it had the same feature value across all exemplars of both categories and, therefore, was irrelevant to classification (i.e., the Irrelevant dimension). Figure 1B shows the category structure of the stimuli used in the task. This stimulus structure is important because selective attention should result in optimizing attention of the deterministic dimension,

and subsequent learning of only this dimension. In contrast, distributed attention should result in learning of some or most of the dimensions.

The stimuli were organized into two sets for the two phases of the experiment. One set was presented to the participants during Phase 1 and its complement set was presented during Phase 2. The two sets were identical, except that the Deterministic and Irrelevant dimensions swapped roles (i.e., Deterministic dimensions in one phase became Irrelevant in the other, and vice versa). The Probabilistic dimensions and the category labels remained the same in both phases. After participants learned one set of stimuli during training in Phase 1 of the experiment and tested on this set, the stimuli were unexpectedly replaced with the complementary set for Phase 2. The order of the sets (Phase 1 and Phase 2) was counterbalanced between participants.

There were two pairs of these stimulus sets that were counterbalanced between participants. The pairs were similar, with two important differences. First, the dimensions that were Deterministic/Irrelevant differed between the pairs. For one pair, hands and feet were Deterministic/Irrelevant dimensions, whereas for the other pair, tail and neck button were Deterministic/Irrelevant dimensions. And second, the assignment of Probabilistic features to the categories differed between the two pairs.

During test sessions, two types of items were presented to participants. The first type, the “Match” items were identical to the stimuli presented during training. The second type, the “Conflict” items, were hybrid stimuli that possessed Probabilistic features that predicted one category label and the Deterministic feature that predicted the opposite category label (see Figure 1B). These items allowed us to determine whether participants’ category judgments were based on the single Deterministic dimension or on multiple Probabilistic dimensions. Responses consistent with the Probabilistic dimensions suggest a similarity-based categorization strategy (which

requires a distributed pattern of attention), whereas responses consistent with the Deterministic dimension indicate selective attention. Analysis of Conflict items is particularly important in Phase 2, because responding based on the new Deterministic dimension requires attending to and learning from it—which would be severely impaired if learned inattention had occurred.

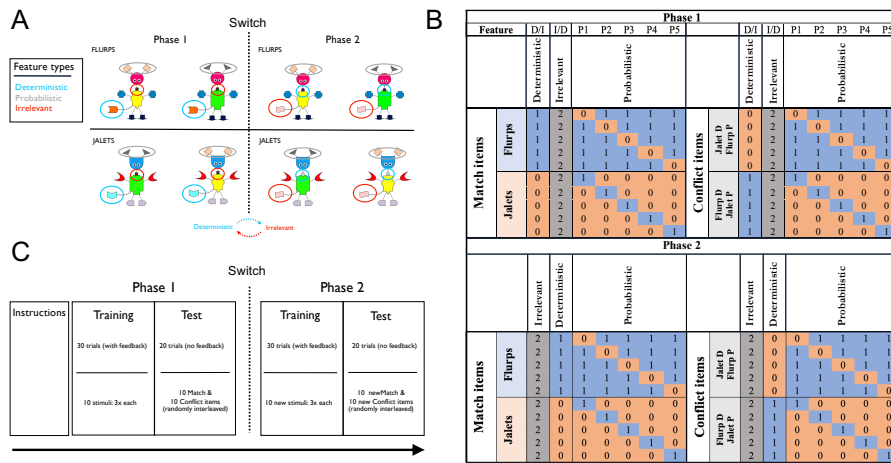


Figure 1. Experiment Design. (A) Example stimuli used the task, where Flurps are shown on the top row and Jalets are shown on the bottom row. In this example, the tail and neck button are Deterministic and Irrelevant features in Phase 1, respectively, and swap roles in the Phase 2. (B) Category structure of Flurps and Jalets by feature value, in Phase 1 (top) and Phase 2 (bottom). Match items (left) were seen in both training and testing sections, while Conflict items (right) were seen only during testing. “P1-P5” represents the five Probabilistic dimensions. “D/I” is the dimension that was Deterministic in Phase 1 and Irrelevant in Phase 2. “I/D” is the dimension that was Irrelevant in Phase 1 and Deterministic in Phase 2. (C) Experimental design with number of trials indicated per each experimental period.

Procedure

Participants completed a classification learning task, while their eye gaze was tracked. The experiment was divided into two phases (Figure 1C). Both phases contained a training section (30 trials with feedback) followed by a testing section (20 trials with no feedback). Following instructions, participants learned to classify two categories of novel creatures in Phase 1, and then in Phase 2 an unannounced switch occurred wherein the previously Deterministic dimension and the previously Irrelevant dimension swapped roles. The formerly Deterministic dimension became Irrelevant: it took on a new, previously unseen, value that was fixed across all stimuli of both categories, providing no information about the correct category label. After the switch, the formerly Irrelevant dimension became Deterministic: it took on two new values that perfectly predicted category membership in Phase 2 (see Figure 1B). Participants were given no warning that this switch would occur. For the sake of clarity, we use the terms D/I for the dimension that was Deterministic in Phase 1 and Irrelevant in Phase 2, and I/D for the dimension that was Irrelevant in Phase 1 and Deterministic in Phase 2. Each of the two phases of the experiment lasted approximately 8 minutes, and the entire task (including eye-tracker set-up) took about 20 minutes.

Instructions

At the beginning of the experiment participants were informed that they would see two kinds of creatures called Flurps and Jalets, and that they needed to figure out which ones were which. They were then given information about the Deterministic and Probabilistic dimensions. For Probabilistic dimensions they were told that *most* of the members of the category had a particular feature value on that dimension, while being shown that feature in isolation. For the Deterministic dimension participants were shown both feature values and told that *all* members of one category (e.g., Flurps) had one feature value, whereas *all* members of the contrasting category

Commented [SV1]: This is a discrepancy with D features' description. Were they shown one value here or 2 (as was the case for the D features).

Commented [NB2R1]: They were shown one at a time for the P features

(e.g., Jalets) had another feature value. The Irrelevant dimension was never mentioned in the instructions. Participants were given the chance to ask questions after the instruction period before starting the task. The experimenter ensured that the participant understood the task before the first trial began, and let the participant know that they could ask questions at any time during the study.

Training

Training in each phase consisted of 30 trials. In each block of 10 trials, the ten training exemplars (the Match items), five from each category, were presented in random order. Therefore, participants saw each exemplar three times throughout training. On each training trial, one stimulus was presented in the middle of the screen, and participants indicated whether they thought it was a Flurp or a Jalet. Adults made responses by pressing a button on a controller, whereas children made their response verbally, and the experimenter recorded the response by pressing a key on a keyboard. The stimulus remained on the screen until a response was made. After each response, corrective feedback was then given, attempting to attract attention to the overall appearance and to the Deterministic dimension. For example, in the case where hands were the Deterministic dimension, feedback would be “Correct this is a Flurp. It looks like a Flurp and has the Flurp hands”, or “Oops this is actually a Jalet. It looks like a Jalet and has the Jalet hands.” In Phase 2, after the unannounced switch, feedback was simplified to mention only the correct category without drawing attention to the dimensions (e.g., “Correct this is a Flurp.”). While this change in feedback may have given participants some indication that a change had occurred, it was necessary so that participants would need to figure out on their own the new contingencies between dimensions and categories. The stimulus remained on the screen during feedback until a button was pressed to begin the next trial.

Testing

Testing in each phase consisted of 20 trials: 10 Match and 10 Conflict items were each presented once, in random order. Similar to training, participants saw the stimuli one at a time and were instructed to indicate whether they thought each creature was a Flurp or Jalet, but no feedback was provided during testing.

Eye Tracking Methods

Participants' gaze was tracked using an EyeLink 1000 eye tracker (SR research, Ontario, Canada) on a hydraulic arm mount with built in speakers. The eye tracker measured eye gaze by computing pupil-corneal reflection at a sampling rate of 250 Hz monocular, with accuracy averaging 0.5°. The eye-tracking device was located inside a darkened testing room, enclosed by curtains. Participants were centrally positioned approximately 60 cm from the eye tracker and the 1280x1024 pixel display monitor. After initial setup, 5-point calibration was conducted prior to the start of the main task (i.e., prior to Phase 1 training). Calibration was repeated at the start of each new section (i.e., Phase 1 test, Phase 2 training, Phase 2 test). During the task, drift correction was conducted every five trials to ensure accurate calibration throughout the experiment.

For the analyses, eight areas of interest (AOIs) of the stimuli were defined as rectangular areas surrounding each feature. These AOIs varied in size from 2° by 2° at the smallest (the neck button) to 4.3° by 4.7° at the largest (the body). There were two AOIs for the hands because there was one hand on each side of the body. The two hand AOIs were treated as a single AOI, resulting in seven AOIs (one per dimension) used for the analyses. Timepoints in which the participant's gaze was not within any of the stimulus AOIs were excluded from subsequent analyses. On average, 36.0% of participants' eye-tracking datapoints were excluded for this reason. In this way, we only analyzed points in which participants were looking at the stimulus, as our analyses concentrated on which and how many stimulus features were attended. Across all participants, 16

(out of 2040) trials were excluded for having no remaining useable datapoints. For each trial, we aggregated the total time that the participant's gaze was within each AOI and calculated the proportion of time spent looking at each dimension. Analyses were conducted on these proportions on the basis of the dimension's role (i.e., Deterministic, Irrelevant, or Probabilistic). For the Probabilistic dimensions, the mean was calculated to facilitate direct comparison to the other dimension roles. We focused our analysis on the fixation proportions for each feature (rather than raw looking times) because we were primarily interested in the extent to which participants' attention was distributed across multiple dimensions compared to selectively focused on one dimension. Eye-tracking data were processed and analyzed using custom code written in Python and R.

Results

Behavioral Results

We first analyzed behavioral measures to assess how well participants learned and what drove their categorization decisions, both before and after the unannounced shift. Figure 2 shows (by age and Phase) (A) categorization accuracy during training, (B) Match item accuracy during testing, and (C) responses to Conflict items during testing. One noticeable deviation from previously observed results was that children exhibited lower accuracy during Phase 2 training—59.9% correct here vs. 64.5% in children of the same age group in Blanco and Sloutsky (2019), and their responses were near chance across both item types during the test section of Phase 2 (57.8% accuracy for Match Items and 53.8% deterministic responses for Conflict Items). One possible explanation for this pattern of results is that children may be exhibiting fatigue during this study, which might be due to the inclusion of eye-tracking that required equipment set up and calibration. In light of this possibility, we identified a subset of children who learned well in Phase

2, and we performed exploratory analyses on this group to ensure that the main findings were not a result of fatigue (or to identify any effects that result from fatigue). If these children's responses and attentional patterns appear similar to adults', there may be reason to believe that the observed developmental effects in the study are due to fatigue effects. We classified child participants as "Good Learners" ($N = 10$) if they achieved an accuracy of 0.7 or higher on Match items during the test section of Phase 2. Because Phase 2 test was the last part of the experiment, participants who were still performing well at that point were unlikely to have experienced substantial fatigue. The Phase 2 analyses below are repeated on the subsample of Good Learners in a separate subsection at the end of the Results section. Overall, Good Learners were similar to the whole sample of all children (except that performance was higher in Phase 2), and their results further support our main findings.

Phase 1

Learning in Phase 1 was good for all participants and similar for both groups (Figure 2). A 2 X 3 (Age group X 10-trial block) ANOVA found a main effect of block, $F(1,134) = 14.88, p < 0.001, \text{partial-}\eta^2 = 0.10$, and a main effect of age group, $F(1,66) = 23.57, p < 0.001, \text{partial-}\eta^2 = 0.26$, but no interaction, $F(1,134) = 0.21, p = 0.648, \text{partial-}\eta^2 = 0.002$. This shows that both groups improved over time, and adults performed better than children.

During the test in Phase 1, accuracy for Match items was also higher for adults ($M = 93.9\%$) than for children ($M = 73.0\%$), $t(66) = 4.33, p < 0.001, 95\% CI [0.113, 0.306], d = 1.06$. Similarly, responses to the Conflict items differed between the groups, with adults ($M = 88.4\%$) responding based on the Deterministic dimension more often than children ($M = 67.7\%$), $t(66) = 3.30, p = 0.002, 95\% CI [0.082, 0.333], d = 0.81$, indicating that adults were more likely rely on only the single Deterministic feature to make categorization judgments than children. This developmental

difference aligns with findings from previous studies and has been interpreted as evidence that adults attend more selectively than children during learning (e.g., Deng & Sloutsky, 2015a, 2016).

Phase 2

Both groups exhibited an initial cost of the unexpected switch, dropping in accuracy from the end of Phase 1 to the beginning of Phase 2 (see Figure 2). Similarly to Phase 1, a 2 X 3 (Age group X 10-trial block) ANOVA on accuracy during training found a main effect of block, $F(1,134) = 8.01, p = 0.005, \text{partial-}\eta^2 = 0.06$, and a main effect of age group, $F(1,66) = 29.26, p < 0.001, \text{partial-}\eta^2 = 0.31$, but no interaction, $F(1,134) = 0.63, p = 0.430, \text{partial-}\eta^2 = 0.005$.

Children's performance was relatively low during the Phase 2 test for Match items ($M = 57.8\%$), and performance was higher for adults ($M = 81.3\%$) than for children, $t(66) = 4.48, p < 0.001, 95\% \text{ CI } [0.130, 0.340], d = 1.09$. Responses to the Conflict items differed between the groups, with children ($M = 53.8\%$) responding based on the newly Deterministic (I/D) dimension *more often* than adults ($M = 36.6\%$), $t(66) = 2.40, p = 0.019, 95\% \text{ CI } [0.029, 0.316], d = 0.59$. Thus, adults exhibited lower Deterministic responding to Conflict items in Phase 2 compared to children, despite responding overwhelmingly based on the Deterministic dimension in Phase 1. These results indicate that adults tended to rely on one or more of the Probabilistic dimensions to categorize during Phase 2, rather than the new Deterministic (I/D) dimension. This reversal from Deterministic responding compared to Phase 1 in adults suggests that adults may have experienced considerable learned inattention, whereas children did not, which replicates previous findings (Blanco & Sloutsky, 2019). Eye-tracking analyses (presented below) directly confirm this suggestion and extend these findings.

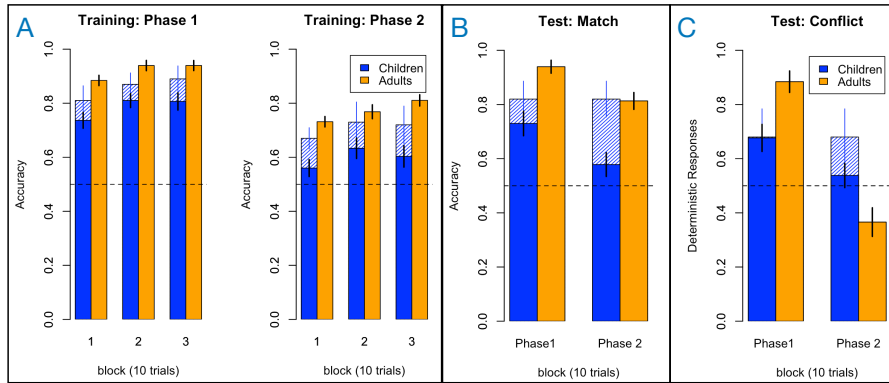


Figure 2. Behavioral Results. (A) Mean accuracy across 10-trial blocks of training, both before and after the switch. (B) Mean categorization accuracy for Match items during the test section of each phase. (C) Mean proportion of responses to Conflict items during the test that were consistent with the Deterministic feature. The solid bars represent the means for all learners. The solid and the striped bars together represent the means for the subsample of children classified as Good Learners. Error bars represent standard errors of the mean.

Patterns of Attention Allocation

As described above, AOIs were defined as rectangles that surround each feature. At each timepoint, we determined which, if any, of the AOIs the participants' gaze fell within. Timepoints in which the participant's gaze was not within any of the stimulus AOIs were removed from subsequent analysis. For each trial the proportion of the trial that the participants' gaze was within each of the AOIs was calculated. We then analyzed these proportions by dimension role (i.e., Deterministic, Irrelevant, or Probabilistic). For the five Probabilistic dimensions, the mean was calculated for easier comparison to the other dimensions.

Figure 3 presents gaze allocation by feature during training for both groups. The pattern in Phase 1 shows that while both groups prioritized looking at the Deterministic (D/I) dimension to some extent, adults showed a greater preference for it than children. Across trials, adults fixated on the Deterministic (D/I) dimension an average of 42.8% of the time in Phase 1, whereas children fixated on it significantly less often—only 24.4% of the time, $t(66) = 2.97$, $p = 0.004$, 95% *CI* [0.060, 0.308], $d = 0.72$. Adults also fixated on the Irrelevant (I/D) dimension ($M = 2.0\%$) significantly less often than children ($M = 6.4\%$), $t(66) = 2.76$, $p = 0.007$, 95% *CI* [0.012, 0.076], $d = 0.67$.

Additionally, the proportion of time spent fixating the Deterministic (D/I) dimension remained relatively stable for children across the training session, while adults continued to optimize their attention by constantly increasing the proportion of time that they were fixating the Deterministic compared to other dimensions (Figure 4). We calculated each participants' slope of the proportion of time fixating the Deterministic (D/I) feature and the Irrelevant (I/D) feature across trials to compare how the two age groups adjusted their attention over time. Adults increased fixations to the Deterministic (D/I) dimension more over time (i.e., had greater positive slopes) than children, $t(66) = 3.64$, $p < 0.001$, 95% *CI* [0.003, 0.011], $d = 0.89$. Adults also decreased fixations to the Irrelevant (I/D) feature across trials more (i.e., had slopes that were more negative) than children, $t(66) = 2.34$, $p = 0.022$, 95% *CI* [0.0002, 0.002], $d = 0.57$.

After the switch, adults showed an interesting pattern in Phase 2. Within a few trials they learned to deprioritize the previously Deterministic, now Irrelevant, (D/I) dimension, but instead of discovering and attending to the newly Deterministic (I/D) dimension, they continued to ignore it (Figures 3 and 4). Instead, they increased their fixations to the Probabilistic dimensions. Furthermore, as shown in Figure 3, this pattern characterized the overwhelming majority of adults:

there were only 3 adults who exhibited any measurable fixation proportion at the I/D feature in Phase 2. This pattern directly demonstrates learned inattention because attending to the dimension that was previously learned to be irrelevant did not increase, despite it providing perfect prediction of the category label in Phase 2. Additionally, one might expect that the newly Deterministic (I/D) dimension having two novel feature values (not seen during Phase 1) would typically draw attention to that dimension—but it did not, further highlighting the power of learned inattention.

Children, conversely, continued to distribute their attention broadly, fixating to all dimensions. Children fixated to the new Deterministic (I/D) dimension more often than adults in Phase 2 (13.8% of the time for children, 5.1% of the time for adults), $t(66) = 3.43, p = 0.001, 95\% CI [0.036, 0.138], d = 0.84$. Children ($M = 14.9\%$) also fixated to the newly Irrelevant (D/I) feature more often than adults ($M = 7.4\%$) in Phase 2, $t(66) = 3.11, p = 0.003, 95\% CI [0.037, 0.124], d = 0.76$. Furthermore, as shown in Figure 3, this pattern of distributed attention was characteristic for the majority of children.

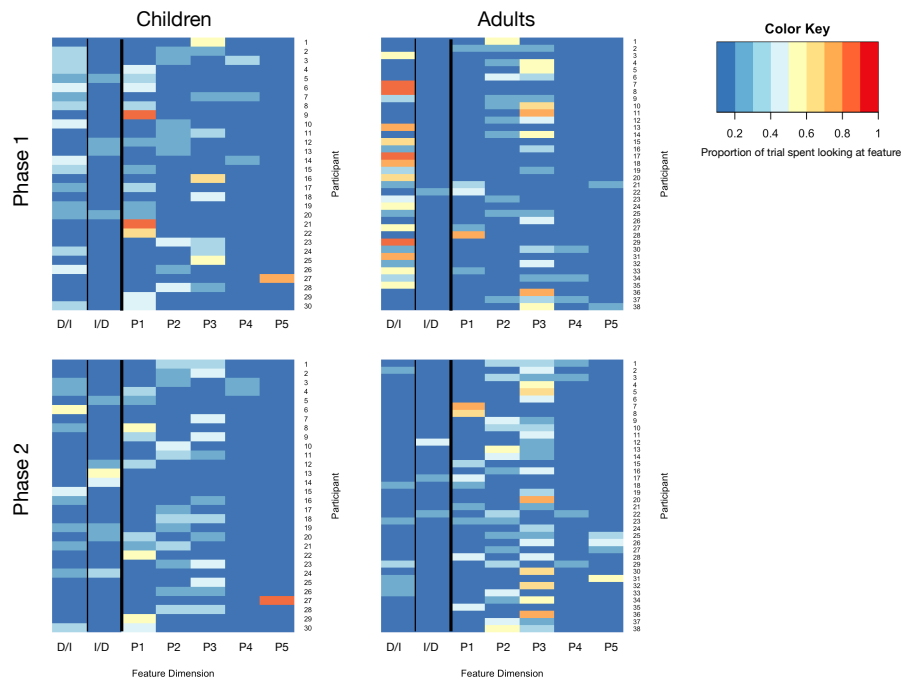


Figure 3. Individual participants' attention allocation. Gaze allocation by stimulus dimensions averaged across training trials is plotted for each age group and phase. "P1-P5" represent the five Probabilistic dimensions. "D/I" is the dimension that was Deterministic in Phase 1 and Irrelevant in Phase 2. "I/D" is the dimension that was Irrelevant in Phase 1 and Deterministic in Phase 2. Note that because there were two stimulus set pairs counterbalanced across participants, the notation here (D, I, and P1-P5) refers to features roles, but not necessarily to feature identity that may differ across the sets (e.g., D feature could be hands for one set in the pair and it could be the tails in the other set).

Commented [NB3]: Maybe it's worth noting here that they are not necessarily the same 5 p dimensions in the same order for all participants—since different participants had different stimuli sets. We could also mention adding it in the letter in response to the one reviewer's comment

Commented [SV4R3]: Done — check if you agree.

Commented [NB5R3]: Looks good

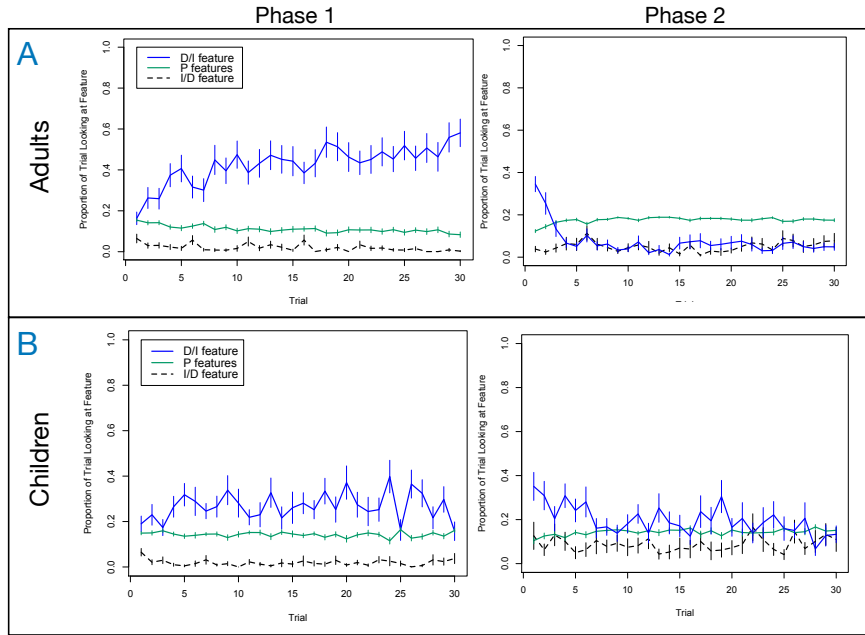


Figure 4. Attention allocation across training trials. (A) Adults' proportion of looking at features by feature role. (B) Children's proportion of looking at feature by role. The line representing the Probabilistic dimensions is based on the average of the five dimensions. Error bars represent standard errors of the mean.

The Benefits of Control: Attention Predicts Learning

In this section, we assess the extent that attentional control provides benefits to learning by examining whether the attentional patterns of individual participants during Phase 1 predict how well they learned (indicated by Match accuracy during Phase 1 test). We also assess whether attentional patterns influence what category representation participants form—whether representations are based on the single Deterministic dimension or on overall similarity (indicated by responses to Conflict items during Phase 1 test). To summarize the extent that each participants'

attention allocation was distributed (as opposed to selective), we calculated the *entropy* of the proportions of looking at the different dimensions. According to information theory, entropy is a measure of diversity or variety among a set of probabilities or proportions (Shannon & Weaver, 1949). Therefore, the entropy of the proportion of looking at different stimulus dimensions encodes the extent that attention allocation was distributed. For each participant, we first calculated the proportion of time spent looking at each feature on each trial. Entropy was then calculated via the following equation,

$$H_t = - \sum_{d=1}^D \log(t_d) t_d \quad (1)$$

where H_t is the entropy of the vector of looking proportions t , and t_d is the proportion of time spent looking at each dimension d on that trial. According to this equation, if a participant looked selectively at only a single feature for an entire trial, entropy would approach 0, whereas maximum entropy would be achieved if they distributed attention equally to all seven dimensions. For the analyses and plots below, entropy was normalized (by dividing it by the maximum possible entropy for seven dimensions) to be between 0 and 1.

We first compared entropy during Phase 1 training between children and adults (Figure 5). A 2 X 3 (Age group X 10-trial block) ANOVA on entropy during Phase 1 training found a main effect of block, $F(1,134) = 33.90, p < 0.001, \text{partial-}\eta^2 = 0.20$, and a main effect of age group, $F(1,66) = 15.15, p < 0.001, \text{partial-}\eta^2 = 0.19$, and a significant interaction, $F(1,134) = 11.27, p = 0.001, \text{partial-}\eta^2 = 0.08$. Adults ($M = 0.594$) and children ($M = 0.642$) had similar levels of entropy during the first 10-trial block of training, $t(66) = 1.39, p = 0.170, 95\% CI [-0.118, 0.021], d = 0.34$, but by the end of training (the last 10-trial block), adults ($M = 0.442$) had substantially lower entropy than children ($M = 0.606$), $t(66) = 4.20, p < 0.001, 95\% CI [0.086, 0.243], d = 1.03$.

Both groups significantly reduced in entropy over time, as indicated by a significant negative slope of entropy across trials during training: Children: $t(29) = 4.58, p < 0.001, 95\% CI [-0.002, -0.005], d = 0.84$; Adults, $t(37) = 7.03, p < 0.001, 95\% CI [-0.009, -0.005], d = 1.14$. Decrease in entropy was greater in adults than in children, $t(66) = 2.89, p = 0.005, 95\% CI [0.001, 0.006], d = 0.71$. These findings indicate that adults optimized their attention more than children by focusing their attention increasingly more selectively as the experiment went on. Although both groups started with similar attention distribution, adults were substantially more selective by the end of training.

We then assessed how attention distribution was related to responses to Match and Conflict items during the test. We first performed a logistic regression across all subjects predicting accuracy on Match items during Phase 1 test from entropy of gaze allocation during Phase 1 training. Entropy significantly predicted Match item accuracy, $\beta = -0.34, z = -2.819, p = 0.005$. The negative relationship indicates that smaller entropy (i.e., greater selectivity) was related to improved accuracy. We then performed a logistic regression predicting Deterministic responses to Conflict items during Phase 1 test from entropy. Entropy significantly predicted Deterministic responses, $\beta = -0.22, z = -2.189, p = 0.029$. The negative relationship between entropy and Deterministic responses to Conflict trials suggests that selective attention is associated with making category judgments based on the single Deterministic dimension, whereas more distributed attention is associated with relying on overall similarity to make category judgments.

Commented [NB6]: I guess this is where reviewer 3 is asking that age be included in the model. The reasoning here to not include it is that we want to see if entropy alone, without age, could explain subsequent behavior. Maybe we should just state that directly around here somewhere

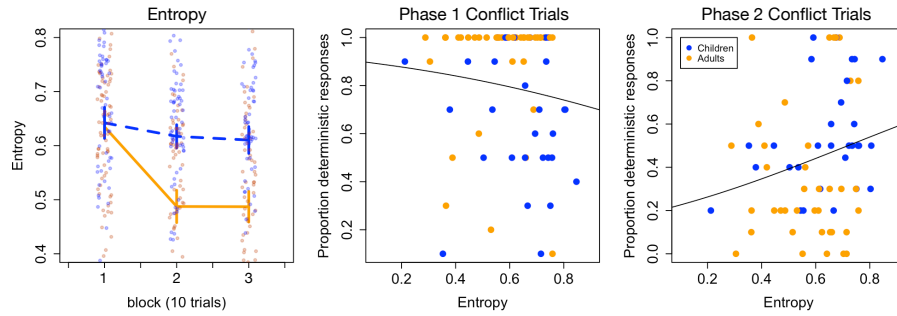


Figure 5. Entropy. (A) Entropy of gaze allocation to different features during Phase 1 training. Blue dots represent Children and orange dots denote adults. Entropy for adults dropped quickly and remained low. Error bars represent standard errors of the mean. (B) Entropy during Phase 1 training was *negatively* related to Deterministic responses to Conflict items during Phase 1. (C) Phase 1 entropy was *positively* related to Deterministic responses to Conflict items during Phase 2. The lines represent the best-fit lines of logistic regressions.

The Costs of Control: Attentional Selectivity Predicts Learned Inattention

We then assessed the costs of control by examining whether selective attention during Phase 1 predicted learned inattention in Phase 2. Because learned inattention may be the result of selective attention (low entropy) or of the process of attention optimization (negative entropy slope) we performed a logistic regression predicting responses to Phase 2 Conflict items from Phase 1 entropy during training, Phase 1 entropy slope (across training trials), and their interaction. Across all participants there was a significant interaction, $\beta = -0.34$, $z = -3.83$, $p < 0.001$. This relationship also held within the group of adults, $\beta = -0.58$, $z = -4.60$, $p < 0.001$, but not within the group of children, $\beta = 0.07$, $z = 0.46$, $p = 0.643$.

We then examined both predictors separately to get a better understanding of this interaction. A logistic regression predicted responses to Conflict items on the test in Phase 2 from entropy in Phase 1, $\beta = 0.28$, $z = 3.53$, $p < 0.001$. This positive relationship suggests that lower entropy (more selectivity) in Phase 1 predicts lower Deterministic responding (more learned inattention) in Phase 2. This same relationship held with the group of children, $\beta = 0.34$, $z = 2.79$, $p = 0.005$, but was not significant within adults, $\beta = 0.09$, $z = 0.87$, $p = 0.387$. By contrast, entropy slopes did not predict learned inattention, $\beta = 0.07$, $z = 0.94$, $p = 0.346$. These results indicate that across the age groups, and for children, higher entropy (more distributed attention) was directly associated with less learned inattention.

Good Learners

In this section, we repeat Phase 2 analyses above using the subsample of children ($N = 10$) classified as Good Learners rather than the entire sample of children.¹ The block by age group ANOVA on Phase 2 training accuracy showed that there was a main effect of block on training accuracy, $F(1,94) = 8.09$, $p = 0.005$, $partial-\eta^2 = 0.08$, but no significant effect of age group, $F(1,46) = 2.62$, $p = 0.113$, $partial-\eta^2 = 0.05$, and no interaction, $F(1,94) = 0.21$, $p = 0.648$, $partial-\eta^2 = 0.002$. The only deviation from the full sample is that Good Learners performed as well as adults. Good Learners ($M = 86.0\%$) did not differ from adults ($M = 81.3\%$) in Match accuracy during Phase 2, $t(46) = 0.72$, $p = 0.47$, $95\% CI [-0.083, 0.176]$, $d = 0.26$, which is not surprising considering that Good Learners were identified by having high accuracy on these items. Crucially, like the full sample, Good Learners classified the Conflict items in Phase 2 on the basis of the

¹ Phase 1 results are not reported here since there is no evidence of fatigue in Phase 1, but overall, results for Good Learners from Phase 1 closely match the results for the full sample.

Deterministic (I/D) feature *more* often than adults ($M_{\text{Good Learners}} = 68.0\%$ vs. $M_{\text{Adults}} = 36.6\%$), $t(46) = 2.78$, $p = 0.008$, $95\% \text{ CI } [0.086, 0.542]$, $d = 0.99$.

Importantly, in Phase 2, Good Learners fixated the new Deterministic (I/D) feature more often than adults, $t(46) = 3.70$, $p < 0.001$, $95\% \text{ CI } [0.055, 0.188]$, $d = 1.31$, similar to the full sample. Critically, similar to the full sample of children, Phase 1 entropy significantly predicted Deterministic responses to Conflict items in Phase 2, $\beta = 0.75$, $z = 3.14$, $p = 0.002$.

These analyses confirm that our main findings are not an artifact of children experiencing fatigue at the end of the experiment. In fact, whereas Good Learners performed as well as adults in Phase 1 and on Match items during Phase 2 test, their responses to Conflict were even more Deterministic than the full sample (and were hence even more unlike adults). Therefore, even children who learned and performed as well as adults, experienced much less learned inattention than adults in Phase 2.

Summary of Results

Overall, both groups learned well in Phase 1, but adults optimized their attention more during learning compared to children. They increased their looking to the Deterministic (D/I) dimension and decreased it to the Irrelevant (I/D) dimension more than children, leading to overall greater attentional selectivity (indicated by lower overall entropy). Adults also overwhelmingly classified Conflict items based on the Deterministic (D/I) dimension in Phase 1. After the switch (Phase 2), adults stopped fixating the previously Deterministic, now Irrelevant, (D/I) dimension and continued ignoring the previously Irrelevant, now Deterministic (I/D) dimension—demonstrating substantial learned inattention. Naturally, they also did not use the new Deterministic (I/D) dimension to classify Conflict items in Phase 2—a complete reversal from Phase 1. Children, conversely, fixated on all dimensions approximately equally in Phase 2, and

were much more likely than adults to classify based on the new Deterministic (I/D) dimension. Most importantly, attention allocation during learning in Phase 1 predicted these responses to Conflict items in Phase 2 on an individual participant basis. More selective attention in Phase 1 led to better learning, but also to more learned inattention. Learned inattention, in turn resulted in a failure to discover new Deterministic features in Phase 2 and reliance on Probabilistic features when responding to Conflict items.

Discussion

This study demonstrates that differences in cognitive control between adults and children have important implications which sometimes benefit children. In our study these differences in cognitive control manifested as differences in how attention was allocated (and optimized) during category learning. Whereas previous studies have shown that children sometimes notice things that adults miss which can be advantageous under the right circumstances (Plebanek & Sloutsky, 2017; Blanco & Sloutsky, 2019), no previous study had directly investigated the mechanism by which this advantage arises. The current study directly ties this advantage to broad distribution of attention, which is one of the main consequences of immature attentional control. Children were less likely than adults to attend selectively, but they were also more likely to discover and use a newly diagnostic source of information. Further, this effect was seen not only on the group level, but on the individual level as well. Across participants the level of attention selectivity also predicted a tendency to miss a newly relevant source of information.

Whereas cognitive control is generally considered universally positive, the results of the current study suggest that its benefit may be context dependent. Deploying attentional control to attend selectively involves a tradeoff. While controlling attention to selectively focus on the most relevant sources of information can lead to efficient learning, it also presents substantial risks. In

scenarios where the environment is highly dynamic, and the relevance of sources of information can change, reliance on selective attention may be detrimental, as potentially important sources of information could be missed. In these cases, it may be better to hedge one's bets by distributing attention widely.

Distributed attention appears to be a developmental default in that selective attention requires a greater level of cognitive control than distributed attention. This idea is also supported by corresponding evidence coming from aging research. Older adults, who are known to experience decline in inhibitory control processes, are less able to maintain selective attention to relevant information and are more likely to process task-irrelevant information (Healey, Campbell, & Hasher, 2008). Specifically, older adults, while showing greater detrimental effects of distractors on learning and memory, also demonstrate better memory for distractors (Rowe, Valderrama, & Hasher, 2006), better learning of statistical regularities that involve distractors (Campbell et al, 2012), and greater facilitative effects of priming from distractors (Amer & Hasher, 2014; Kim, Hasher, & Zacks, 2007).

On the basis of the reported and reviewed evidence, we suggest an alternative, complimentary, view on the maturity of cognitive control: that protracted immaturity of cognitive control in children may serve an adaptive function (see Liquin & Gopnik, 2022 for related evidence, as well as Bjorklund, 1997; Bjorklund & Green, 1992; Bruner, 1972, for more general benefits of immaturity). As children undergo development, their environment, their goals, and their construal of what is relevant all change rapidly, thus making the world around them volatile. Additionally, children often do not know what is relevant in a given situation and how to identify it. Increasingly with development, children become more and more prepared to identify what is relevant to attend to it, given the goals and task demands. Distributed attention stemming from

immature cognitive control may serve to protect children from the potential costs of selective attention until they have accumulated the knowledge necessary for effective control. Without that experience they may use selective attention in a way that is misguided, engage it prematurely, and fall into learning traps.

In fact, there is some evidence that distributed attention may facilitate learning in infants. For example, Deng and Sloutsky (2015b) demonstrated that greater distribution of attention (as measured by the number of gaze shifts) was accompanied by more robust category learning in 8-to-12-month-olds. Similarly, Jankowski et al. (2001) demonstrated that inducing distributed attention in 5-month-old infants (by presenting a dynamic visual cue in various parts of the screen) during attention training resulted in a more distributed pattern of attention (compared with no training) and in more efficient learning of the presented items. Therefore, distributed attention may be beneficial to children's learning early in life—perhaps because they would be otherwise more prone to traps than adults due to having less experience.

Limitations and Future Directions

This study has a number of potential limitations. First, as previously noted, the study may have approached the limit of sustained engagement for children of this age group, leading to fatigue and compromising the quality of the data collected toward the end of the experiment. Although our exploratory analysis of “Good Learners” supports our main results and suggests that it is unlikely that our main findings were the result of fatigue, fatigue may nonetheless introduce noise into the data, potentially obscuring more subtle effects. A further limitation is that the experiment was not designed to make use of response time data, which, in theory, would provide further insight into the cognitive processes of children and adults. Response modes were different between children and adults—adults pressed a button themselves, while children verbally responded, and

the experimenter pressed a button to record the response—which made measurement of children’s response times too imprecise for proper comparison. Another limitation is that only a single age group of children was included. Future work with a wider range of ages will be needed to better understand the developmental trajectory of these process as cognitive control develops.

An important avenue of future research will be to firmly establish causal links among immature control, distributed attention, and the ability to avoid learning traps early in development. The current study provides an important first step in this direction by measuring attention directly and relating individual differences in attention allocation to participants’ vulnerability to learning traps. In future studies, causal links could be established through a combination of manipulating attention and longitudinal designs. For example, an attentional load could be introduced to overload control systems. Under these circumstances, we may expect that adults will exhibit attentional patterns and behavioral responses similar to children. Longitudinal studies would allow for the use of cross-lagged panel analysis to better understand the potential causal contingencies between these cognitive processes. For example, if within-subject increases between Timepoint₁ and Timepoint₂ in the ability to attend selectively predict increases in learned inattention at a later time, this finding would support the idea that the links are indeed causal in nature.

Conclusions

The reported study is the first to link attentional control and the possibility to end up in a learning trap. The results suggest that cognitive control may be a double-edged sword: immaturity may actually be adaptive for children such that distributed attention better suits their needs given their level of knowledge and experience. It prevents young children from over-exploiting predictive sources of information and protects them from the potential pitfalls of that exploitation,

instead inducing an exploratory approach to early learning. Although the information around us is vast and it may be beneficial to select some aspects over others, selective attention can lead us down the metaphorical rabbit hole, creating learning traps in which we completely neglect valid sources of information. Distributed attention commands caution, moving slowly to avoid these pitfalls, which may be critically important for children, who have only begun to learn to navigate the vast sea of information around them in the world.

References

- Amer, T., & Hasher, L. (2014). Conceptual processing of distractors by older but not younger adults. *Psychological Science*, 25, 2252-2258.
- Best, C. A., Yim, H., & Sloutsky, V. M. (2013). The cost of selective attention in category learning: Developmental differences between adults and infants. *Journal of Experimental Child Psychology*, 116, 105–119. <http://dx.doi.org/10.1016/j.jecp.2013.05.002>.
- Bjorklund, D. F. & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *American Psychologist*, 47, 46–54. <https://doi.org/10.1037/0003-066x.47.1.46>.
- Bjorklund, D. F. (1997). The role of immaturity in human development. *Psychological Bulletin*, 122, 153–169. <https://doi.org/10.1037/0033-2909.122.2.153>.
- Blanco, N. J. & Sloutsky, V. M. (2019). Adaptive flexibility in category learning? Young children exhibit smaller costs of selective attention than adults. *Developmental Psychology*, 55, 2060–2076. <https://doi.org/10.1037/dev0000777>.
- Bruner, J. S. (1972). Nature and uses of immaturity. *American Psychologist*, 27, 687-708.
- Campbell, K. L., Zimmerman, S., Healey, M. K., Lee, M., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging*, 27(3), 650.
- Deng, W. S., & Sloutsky, V. M. (2015a). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology*, 51, 392–405. <http://dx.doi.org/10.1037/a0038749>
- Deng, W., & Sloutsky, V. M. (2015b). The role of words and dynamic visual features in infant category learning. *Journal of Experimental Child Psychology*, 134, 62-77.
- Deng, W. S., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, 91, 24-62.

- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
<http://dx.doi.org/10.1146/annurev.ne.18.030195.001205>
- Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: the effects of endogenous cueing and of distractors. *Developmental Science*, 4(2), 209–219.
doi:10.1111/1467-7687.00166
- Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science*, 13, 622– 635.
<http://dx.doi.org/10.1111/j.1467-7687.2009.00921.x>
- Healey, M. K., Campbell, K. L., & Hasher, L. (2008). Cognitive aging and increased distractibility: Costs and potential benefits. *Progress in brain research*, 169, 353–363.
- Hoffman, A. B., & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility. *Journal of Experimental Psychology: General*, 139, 319–340. <http://dx.doi.org/10.1037/a0019042>
- Jakobsen, K. V., Frick, J. E., & Simpson, E. A. (2013). Look here! The development of attentional orienting to symbolic cues. *Journal of Cognition and Development*, 14(2), 229–249. doi:10.1080/15248372.2012.666772
- Jankowski, J. J., Rose, S. A., & Feldman, J. F. (2001). Modifying the distribution of attention in infants. *Child Development*, 72, 339–351. <http://dx.doi.org/10.1111/1467-8624.00282>
- Kim, S., Hasher, L., & Zacks, R. T. (2007). Aging and a benefit of distractibility. *Psychonomic bulletin & review*, 14(2), 301–305.

- Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin & Review*, 7, 636–645.
<http://dx.doi.org/10.3758/BF03213001>
- Leclercq, V., & Siéroff, E. (2013). Development of endogenous orienting of attention in school-age children. *Child Neuropsychology*, 19(4), 400–419.
doi:10.1080/09297049.2012.682568
- Liquin, E. G. & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, 218, 104940.
<https://doi.org/10.1016/j.cognition.2021.104940>
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current directions in psychological science*, 21(2), 71–77.
- Pearson, D. A., & Lane, D. M. (1990). Visual Attention Movements: A Developmental Study. *Child Development*, 61(6), 1779–1795. doi:10.1111/j.1467- 8624.1990.tb03565.x
- Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, 28, 723–732.
<http://dx.doi.org/10.1177/0956797617693005>
- Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta psychologica*, 86(2-3), 227–272.
- Rich, A. S. & Gureckis, T. M. (2018). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, 147, 1553–1570. <https://doi.org/10.1037/xge0000466>

- Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 275–285.
<http://dx.doi.org/10.1037/0096-1523.7.2.275>
- Rowe, G., Valderrama, S., Hasher, L., & Lenartowicz, A. (2006). Attentional disregulation: a benefit for implicit memory. *Psychology and aging*, 21(4), 826.
- Schul, R., Townsend, J., & Stiles, J. (2003). The development of attentional orienting during the school-age years. *Developmental Science*, 6(3), 262–272. doi:10.1111/1467-7687.00282
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014). Orienting attention within visual short-term memory: Development and mechanisms. *Child Development*, 85(2), 578–592. doi:10.1111/cdev.12150
- Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences*, 12(9), 342–348. doi:10.1016/j.tics.2008.05.007
- Wainwright, A., & Bryson, S. E. (2005). The development of endogenous orienting: Control over the scope of attention and lateral asymmetries. *Developmental Neuropsychology*, 27(2), 237–255. doi:10.1207/s15326942dn2702_3
- Wong-Kee-You, A. M. B., Tsotsos, J. K., & Adler, S. A. (2019). Development of spatial suppression surrounding the focus of visual attention. *Journal of Vision*, 19(7), 9. doi:10.1167/19.7.9