Attention to Numerosity Varies Across Individuals and Task Contexts

Michèle M. M. Mazzocco, Jenny Yun-Chen Chan, Jeffrey K. Bye

Emily R. Padrutt, Taylor L. Praus-Singh, Sarah Lukowski,

Ethan C. Brown, Rachel E. Olson

University of Minnesota

(2020, In Press, *Mathematical Thinking and Learning*)

Author Note

Michèle Mazzocco, Institute of Child Development, University of Minnesota; Jenny Yun-Chen Chan, Institute of Child Development, University of Minnesota; Jeffrey K. Bye, Institute of Child Development, Department of Educational Psychology, and Research Methodology Consulting Center, University of Minnesota; Emily R. Padrutt, Institute of Child Development, University of Minnesota; Taylor Praus-Singh, Institute of Child Development, University of Minnesota; Sarah Lukowski, Institute of Child Development, University of Minnesota; Ethan Brown, Department of Educational Psychology and Research Methodology Consulting Center, University of Minnesota; Rachel E. Olson, Institute of Child Development, University of Minnesota.

Jenny Yun-Chen Chan is now at the Department of Social Science and Policy Studies, Worcester Polytechnic Institute.

Acknowledgments. This work was supported by NSF Award #1644748 awarded to the first author. Study 1 was partially supported by the Heising-Simons Foundation Award 2016-078, as part of the Development and Research in Early Mathematics Education (DREME) Network. A portion of Study 1 was completed by the fifth author as partial fulfillment of her Bachelor's degree honors thesis, University of Minnesota. Study 2 was partially supported by a University of Minnesota Grant In Aid Program Award, awarded to the first author. The authors thank the many participants in both studies, and research assistants from the Math and Numeracy Lab at University of Minnesota involved in data collection, scoring, and, entry.

Correspondence concerning this article should be addressed to Michèle Mazzocco, Institute of Child Development, University of Minnesota, Minneapolis, MN 55455 USA. Email: mazzocco@umn.edu

Abstract

Children's spontaneous focus on numerosity (SFON; Hannula & Lehtinen, 2005) is described as an unprompted tendency that is stable across contexts. The attention to number task (AtN), an experimental forced-choice picture-matching task designed to evaluate select aspects of children's focus on numerosity, may reveal whether task materials can implicitly prompt children to focus on numerosity. In two studies, we replicate earlier findings showing an effect of task context on children's performance on the AtN: When asked to identify one or more matches to a target picture from an array of four options, the frequency with which preschoolers and adults identify a numerosity-based match varies as a function of the features on which the remaining match options are based. We addressed a limitation of the original AtN study (Chan & Mazzocco, 2017) by including novel combinations of features as additional trials, with which we continued to demonstrate contextual effects. We also showed that adults seemed more susceptible than children to be primed to attend to numerosity on subsequent trials. Children's focus on numerosity under these experimental conditions was remarkably low. We discuss the implications of these findings for better understanding the SFON construct.

Keywords: numerosity; SFON; contextual effects; early mathematics; attention to number

Attention to Numerosity Varies Across Individuals and Task Contexts

This special issue, and the present study, are devoted to better understanding the nature of children's tendencies to attend to mathematical features, including numerosity. Some researchers define the spontaneous focus on numerosity (SFON) that is observed in children as a stable "non-guided tendency" (Hannula & Lehtinen, 2005, p. 238) to focus on exact numerosity "across time and across different contexts" (Hannula, Mattinen, & Lehtinen, 2005, p. 63). That is, proponents of the SFON construct allege a domain-specific dispositional characteristic rather than general attentional capacities (Hannula, Lepola, & Lehtinen, 2010) or numerical competence (Verschaffel et al., 2018). More recently, however, other researchers have suggested that behavioral responses that occur during SFON research tasks may be subject to implicit prompts linked to task stimuli (Chan & Mazzocco, 2017) or task demands (Batchelor, Inglis, & Gilmore, 2015), even if such prompts are unintended. These dispositional characteristics of SFON and contextual influences on behaviors during SFON tasks do not necessarily contradict each other, yet an understanding of their relative roles may enrich our understanding of SFON as a construct. For example, we (Chan & Mazzocco, 2017) developed the Attention to Number task (AtN) to evaluate attention to numerosity in the context of other features to which children might also attend, such as color or shape. The AtN task differs from traditional SFON tasks in marked ways, because it was designed to experimentally manipulate and isolate one potential task characteristic— the competing perceptual features present in a visual context—that may implicitly prompt individuals to attend to numerosity (or another feature) during a task. In the present study, we use a slight modification of our original AtN task to replicate and expand upon our initial findings, in order to better understand the potential role of task contexts as implicit cues that prompt both children and adults to attend to numerosity.

A better understanding of the SFON construct is timely in view of growing interest in it and related tendencies (e.g., spontaneous focus on quantitative relations; McMullen, Hannula-Sormunen, & Lehtinen, 2013). This interest stems from reports of SFON's predictive relation with later mathematics achievement (e.g. Hannula et al., 2010; Hannula-Sormunen, Lehtinen, & Räsänen, 2015). This relation is meaningful given the wide range of quality of life indicators associated with mathematics achievement levels, ranging from occupational success in science, technology, engineering, and mathematics (STEM) careers to everyday behaviors (e.g., health care decision making, planning schedules, designing projects, managing finances, and following our favorite sports teams; see McCloskey, 2007 for a review). The associations between early mathematical thinking and later mathematics achievement have motivated education researchers to identify early indicators of mathematics difficulties, and have even motivated policy makers and education foundations to promote mathematics learning opportunities beginning in early childhood (e.g., National Governors Association, 2014; Education Endowment Foundation, 2020). Accordingly, it behooves us to better understand SFON tendencies and their potential influence on early mathematics. Towards these goals, in the present study we rely on the AtN task because it may reveal specific circumstances that may constrain or enhance when and whether individuals focus on numerosity, which in turn may have implications for understanding the SFON construct. We note that the AtN task itself – a measure to quantify the specific behavior operationalized as the frequency of attending to numerosity – does *not* represent a novel construct, consistent with the intentions behind the original development of the task (Chan & Mazzocco, 2017).

The general finding that associations exist between early and later mathematics competence are drawn from studies of early numeracy skills (Geary, Hoard, Nugent, & Bailey, 2013; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Mazzocco & Thompson, 2005), and they invoke interest in numeracy as a learning objective, a readiness screening target, or both. A question to emerge from the early numeracy research concerns the extent to which young children's attention to numerosity is foundational for, rather than a reflection of, emerging number competence — regardless of its spontaneity. It follows that related research questions concern whether children's behavior during SFON tasks is an appropriate indicator of mathematics learning readiness, and if promoting SFON in young children is a useful means by which to diminish early mathematics difficulties. When considering such applications of the SFON research, it may be useful to further delineate whether SFON behaviors (i.e., unprompted attention to numerosity in play and other contexts) reflect reliable and stable differences in cognitive processing or dynamic behaviors that reflect children's learning history. The pursuit of these questions may deepen our understanding of the SFON construct by broadening our thinking about what environmental triggers may implicitly promote children's attention to numerosity.

The SFON construct emerged in the early 2000s amidst alternative approaches to measuring or understanding attention to numerosity (Baroody, Li, & Lai, 2008), and it has clearly gained momentum (Verschaffel et al., 2018). More recently, Batchelor and colleagues (2015) demonstrated that SFON tendencies vary with task features or demands, which challenges notions of stability across contexts. Batchelor and colleagues found that children focused on numerosity more often when a task required them to imitate an experimenter placing letters in a postbox compared to a task that required describing pictures that included numerical features. Chan, Praus-Singh, and Mazzocco (2020) found that parents' and children's math-based comments during play occurred more frequently when parent-child dyads interacted with blocks

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

compared to dramatic play props (i.e., a kitchen set), although the frequency of their number-based comments was comparable across these play scenarios. Despite differences in the types of materials used across these two studies, both studies support the notion that the lack of *explicit* prompts in a task does not rule out its *implicit* prompts.

The AtN was developed to experimentally evaluate this notion of implicit prompts by testing children's attention to numerosity relative to other visual features in the task (Chan & Mazzocco, 2017). Briefly, the AtN task requires selecting which of four pictures best matches a target picture, when each of three options matches the target on only one feature, one option does not match the target on any features (i.e., is a foil), and none of the four options matches the target exactly. Among these options, there is no normatively 'correct' response. The match options vary in their numerosity, shape, color, pattern, orientation, or positional location (within the frame of the picture). Per trial, combinations of three features are present among the match options, and the combinations of features themselves are a condition that varies across trials (Figure 1), as follows. On 16 of the 32 AtN test trials, one response option per trial matches the target on set size (i.e., numerosity) and is classified as a number-based match.[1] The other 16 trials lack number-based match options, and instead include a response option that matches the target on rotational orientation. In 8 of each set of 16 trials, the three response options (in addition to the number- or orientation-based match) match the target on shape, color, or none of these features (i.e., a foil; see set A1 and B1 in Figure 1). In the remaining eight trials, the three additional response options match the target on either location or pattern (instead of shape- and

[1] We (Chan & Mazzocco, 2017) named our task the "Attention to number task," and in that task, matches based on number are always tied to set size (numerosity), rather than digits or number words. As such, the term "number" is synonymous with numerosity in descriptions of the AtN task. Still, to be consistent with language conventions used in our original paper, we use the word "number" when describing the AtN task and scores. We use the term numerosity when referring to SFON, to be consistent with the original SFON reports. These terms are used synonymously in this report, unless we refer explicitly to digits or number words.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

color-based matches) or none of these features (A2 and B2 in Figure 1). When creating the original AtN task, we proposed that shape and color are more salient to children compared to location and pattern, and that therefore the two sets of eight trials comprised two conditions representing competing features of relatively high or low salience, respectively (Chan & Mazzocco, 2017). Importantly, this speculation was based on the outcomes of piloting and not on measures of psychophysiological salience per se, and reflected features evident in everyday toys and objects with which children may engage. Of these features, only one (shape) has been empirically demonstrated to conform to a strong categorization bias, albeit among young children engaged in linguistic tasks (Landau, Smith, & Jones, 1988). The relative salience of the other features was something we inferred and validated in the context of our original AtN study (Chan & Mazzocco, 2017). For example, in that study, individuals selected number-based matches less frequently when number was pitted against shape and color (features of high salience) compared to pattern and location (features of low salience). This finding suggests that SFON may be prompted by task stimuli, but it also raises questions about *why* particular co-occurring features influence the frequency of number-based match responses and what the AtN task may reveal about the nature of the SFON construct.

In order to evaluate how performance on the AtN task may further inform our understanding of the SFON construct, we designed the present study to pursue the three following goals: First, we aimed to replicate our original findings (Chan & Mazzocco, 2017) with a different sample of preschoolers and adults, using the AtN task. This attempt at replication is an essential aspect of evaluating the AtN's reliability. Second, to move beyond replication, we addressed a limitation of the original AtN task—that it involved only two sets of features contrasted between trials. Towards this goal, we included two novel combinations of features across an additional set of trials to our sample of preschoolers. Given our earlier finding that salience of competing features affects AtN task performance (Chan & Mazzocco, 2017), it is important to establish the degree to which low and high salience classifications conform to participants' response patterns across different combinations of features. This was particularly true for preschoolers, who made very few number-based matches overall on the original AtN task (Chan & Mazzocco, 2017). Examining novel combinations of features affords an opportunity to assess the salience of number amidst another combination of features. Finally, we further explored the individual differences in adults' AtN performance we reported in our earlier study and their relation to mathematics achievement scores. We chose to pursue this question with adults rather than children because of the larger range of number-based match scores among adults compared to children, as reported previously (Chan & Mazzocco, 2017).

Study 1

Materials and Method

Participants. Participants were preschool children recruited through a participant pool at the University of Minnesota, in the upper midwestern United States. Parents were contacted and invited to participate with their child in a larger study on parent-child shared reading provided that their child was 4 years old, spoke English, and had no known developmental delay. Of 80 parents contacted, only one child did not meet all eligibility criteria. Of the remaining 79 who expressed interest in the study, 49 kept their appointments. Eight participants were excluded from the study because they did not complete the Attention to Number task—the focal measure of this study—and they missed too many AtN trials to allow data imputation. The final sample of 41 child participants ranged in age from 4 years 6 months to 5 years 0 months (*M* = 4 years 10 months, *SD* = 1.4 months) at testing.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

Materials.

Attention to Number (AtN) task. A primary goal of the study was to replicate the results of our earlier study (Chan & Mazzocco, 2017) using our original version of the AtN task (described in detail in the original report). As summarized in the Introduction, the AtN is a picture-matching task designed to assess attention to set size, relative to alternative features. A key component of the AtN task is its experimental manipulation of the "salience" of the alternative features present per trial. (In this context, salience simply refers to the likelihood that a child will attend to or notice the feature, rather than a specific psychophysiological property). In our original AtN study, we repeated AtN administration to child participants 5 months following their initial AtN session, to establish test-retest reliability. Using paired t-tests we found that the frequency of best match responses for all features (and foils) across the two AtN task administrations did not differ (ps > .16), suggesting that the task is a reliable measure of attention to visual features (Chan & Mazzocco, 2017, p. 68).

Task procedure. We followed our original procedure (detailed in Chan & Mazzocco, 2017) for the practice and experimental trials. In keeping with the original AtN task, four practice trials precede the experimental trials, in order to convey that match choices do *not* match the target *exactly* and that the goal of each trial is to identify the "best" match. Here children were asked to answer quickly, and to just choose the best match. Thereafter, on each experimental trial, the experimenter showed the participant a target picture centered on the top half of the page, and four response choices arranged horizontally on the bottom half of the page. Children were asked to "look at the picture at the top of the page," and to "point to the picture (from the bottom row) that is the best match," all while the target and the four response choices remain in view. Although the experimenter did not explicitly ask children to answer quickly on each trial, she did ask children to "just choose which one you think is the best match."

After completing all "best match" trials, the experimenter revisits trials to provide an opportunity for participants to identify number as a matching feature even if they had not selected number as a best match during the initial trials. In the present study, only a select number of trials were revisited due to time constraints and preschoolers' attentional capacities, and in keeping with Chan and Mazzocco (2017). All participants revisited the same set of eight trials from the original AtN task, four of which included number as a potential matching feature. Unlike the best match trials, children were encouraged to identify as many potential matches as possible on each revisit trial.

Task design. Details of the AtN task design appear in Chan and Mazzocco (2017). The original task is comprised of four practice trials and 32 individual forced-choice experimental trials, each of which appears on a single 8 ½ by 11- inch piece of paper. Each experimental trial included a target picture and four response options. The primary outcome of interest in the AtN task is the relative frequency with which children select a number-based match (out of 16 trials) compared to the frequency with which their match selections are based on competing features. A secondary outcome of interest is the relative frequency with which children select an orientation-based match on the remaining 16 trials. Individual trials from these two sets of 16 trials alternate with one another, to diminish the likelihood of children adopting a perseverative response bias (at least on sequential trials), and to provide a baseline comparison for number compared to orientation as a contrasting low salience feature of interest against the same set of competing features. Accordingly, each trial is comprised of four match options: either a *number* (i.e., numerosity) or *orientation* match, two additional match options based on either *shape* and *color* or *pattern* and *location,* and *one foil*.

The combinations of features per set of trials are systematically manipulated so that *number* and *orientation* compete with features classified as either *high salience* (shape and color) or *low salience* (pattern and location) across equivalent sets of trials (Figure 1), but never with each other. As described briefly in the introduction, the resulting two blocks of 16 trials (hereafter Sets A and B) differ in terms of which features compete with number-based match options on a given set of trials. Set A consists of eight trials wherein number-based match options are combined with shape- and color-based match options (subset A1), and eight trials wherein orientation-based match options are combined with pattern- and location-based matches (subset A2). These two subsets of eight trials are interleaved so that none of the features appear on two consecutive trials. In Set B, number and orientation replace each other as response options across the subsets. That is, Set B consists of eight trials wherein number-based match options are combined with pattern- and location-based matches (subset B1), and eight trials wherein orientation-based match options are combined with shape- and color-based matches (subset B2). Similar to Set A, the two subsets of eight trials in Set B are interleaved. Within all four sets of eight trials, the relative position in which specific stimuli features occur is counterbalanced (e.g., number-based matches occur with equal frequency in all four positions, across trials; the same is true for all features, including foils). With regard to the order in which children view these trials, we (Chan & Mazzocco, 2017) previously tested two possible AtN set orders (i.e., Set A first vs. Set B first) and found that they did *not* affect the frequency of preschoolers' number-based selections; therefore, in the present study we used a fixed order design for the original AtN trials, starting with the B2 trials.

In addition to these two original blocks of AtN trials, we created a third novel block of 16 trials (Set C) and presented this set of trials either before or after the entire original AtN task (Sets B and A). For these 16 trials, we used the same stimuli as those used to create Sets A and B, reconfigured to afford within-trial feature comparisons that were impossible in the original AtN study. Specifically, eight trials were combinations of shape-, color-, and pattern-based matches (and a foil); this combination was classified as the high salience combination based on prior findings. The remaining eight trials were combinations of number-, orientation-, and location-based matches (and a foil), classified as the low salience combination (based on the original study findings (Chan & Mazzocco, 2017)). In creating Set C trials, we excluded *pattern* in the novel low salience combination because of the higher frequency with which children (and adults) selected pattern as a best match when it was pitted against location and number, or location and orientation, in our original study (Chan & Mazzocco, 2017). Based on this higher frequency, we included pattern in our novel high salience combination for Set C. Set C thus affords an opportunity to evaluate if number will be selected more frequently than the two competing features that have demonstrated low salience (orientation, location) in the earlier study. This novel combination may be a more robust test of the effect of low salience competing features on attention to number than the original low salience combination we used in our previous study (i.e., number, location, pattern; Chan & Mazzocco, 2017). Finally, in order to test the possibility of an order effect with the new stimuli, half of the children completed best match trials for Set C *before* completing the best-match trials from the original AtN trial blocks (Sets B and A), and the remaining half completed best match trials for Set C *after* completing Sets B and A best match trials.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

Unlike the eight trials revisited from Sets B and A, Set C trials were revisited by only some participants. Up to two revisits occurred for the Set C trials but only if the child had selected foil matches on those trials. Because these two additional revisits were exploratory and did not occur for all participants, they were excluded from analyses.

AtN task administration time was approximately 20 minutes. The primary variables of interest were the *best match score* (the frequency of number-based or orientation-based matches during best match trials only) and the *total match score* (the total frequency of number- or orientation-based matches *ever* made by a participant, across the best match trials and the revisit trials), each out of a maximum possible of 16. Although the AtN task can be used to measure the best match and total match score for any feature included in the task, it was created as a means to measure attention to *number* (i.e., set size, or numerosity) in the context of additional features. (As explained earlier in this section, we also measured orientation matches to contrast with number matches.)

Brief number knowledge assessment. Children completed a set of basic counting tasks to determine their familiarity with numbers up to 4 (the largest set size appearing in number-based match options). These tasks were used to engage children in verbal counting up to 130, enumerating pictured sets of 2 to 10 items, and producing sets of 3 to 19 tokens. Administering the number knowledge tasks required approximately eight minutes. We recorded participants' highest number counted correctly in the count list and the largest set size they consistently enumerated or produced. Children's performance on these tasks was used to verify their number knowledge level, and to explore the relation between their AtN and counting performance.

Procedure. All participants were tested individually by the same female experimenter (TP-S) in a dedicated University-based research office, following a protocol approved by our institutional review board (IRB), in the context of a larger study on parent-child shared reading. Participants completed the AtN task before the number knowledge tasks so that the latter would not prime for number-based responses.

Results

Preliminary descriptive analyses. We first ran descriptive analyses to verify that participants had sufficient number knowledge to evaluate quantities on the AtN. All participants counted aloud to at least 10 without error and had a count-aloud ceiling between 10 and 130 (M = 38.97, SD = 35.24), so all were included in the analyses. All but two children correctly enumerated sets of 4 or more pictured items, and the remaining two children correctly counted sets of 3 pictured items. All but two children correctly produced sets of 5 tokens: when asked to produce sets of 3, 5, 9, and 19, many children correctly produced sets of 9 (n = 29) or even 19 (n = 14), and nearly all children correctly produced sets of 5 (n = 39). The two remaining children correctly *produced* sets of 3. To be consistent with our earlier work (Chan & Mazzocco, 2017) and to maximize the performance range for this task, all 41 of these children were included in the study and the analyses.

Next, we verified that participants selected best matches without significant deliberation, as reported in the original AtN study. Mean time to complete all 48 best match trials was 8.67 minutes (SD = 1.90 minutes), or an average of 10.84 seconds per trial. This per trial speed was comparable to the average rate of 11.8 seconds observed in our original AtN study (Chan & Mazzocco, 2017).

Finally, we replicated the relative frequency with which each individual feature was selected as a best match from the four options, among the 32 original AtN trials (Table 1). Our frequencies were remarkably consistent with those we reported earlier (Chan & Mazzocco, 2017). The children in our study selected shape- or color-based best matches with comparable degrees of frequency, whereas their matches based on pattern occurred more frequently than those based on location. Across trials for the original Sets A and B, the relative frequency of number- and orientation-based matches were comparable with each other and with the rates we reported earlier (Chan & Mazzocco, 2017). Children rarely selected foils as a best match (on ~2 of 32 trials). The low rate of foil selections also showed that children's selections were not random.

Primary analyses: Replication.

Variables of interest. Consistent with the study we aimed to replicate, our primary outcome variable of interest was *attention to number* (i.e., set size) and (for comparison) *orientation*, on the original AtN task (i.e., Sets A and B), measured as the number of trials on which a number-based (or orientation-based) option was selected as the *best match* or across best match trials plus four revisit matches. Our primary independent variable of interest was the two levels of salience manipulated across trials, corresponding to whether the features that competed with number-based (or orientation-based) matches were classified as *high* salience (i.e., shape and color) or *low* salience (i.e., pattern and location). Order was irrelevant to our replication goals, because the children in the present study completed the original AtN trials in a fixed order (Sets B + A). We did, however, test for order effects related to whether Set C (trials with novel combinations of features) occurred before or after Sets B + A (i.e., the original AtN task), to rule out whether the longer task (from Set C preceding the original AtN) influenced children's number-based matches on the original AtN trials. Thus, we used mixed analyses of variance (ANOVAs) for our primary analyses and conducted separate but parallel mixed ANOVAs for the frequency of *number*-based (and, separately, *orientation*-based) best (or total) match scores under high salience versus low salience conditions (as a within-participants repeated measure) and order (Set C first vs. last, as a between-participants variable). We used Bonferroni-corrected pairwise comparisons for all post-hoc tests examining significant interactions.

Children's attention to number and orientation.

Number-based matches. We conducted a 2 (Salience: high vs. low) × 2 (Order: Set C before or after Sets B + A) mixed ANOVA on number-based best match scores, and replicated the significant main effect of Salience found by Chan and Mazzocco (2017), $F(1, 39) = 31.29$, $p < .001$, $\eta_p^2 = .445$. Children made number-based best matches more frequently when the alternative choices were of low salience ($M = 1.49$ out of 8 possible, $SD = 1.61$) compared to high salience ($M = 0.37$ of 8, $SD = 0.80$). There was no significant main effect of Order, and no Salience × Order interaction, ps ≥ .267, η_p^2s ≤ .031, indicating that the main effect of Salience held for number-based best match scores regardless of whether the AtN trials followed or preceded Set C novel trials.

The results held when we replaced the *best match* score with the *total number match score* (on best match trials plus revisit trials). There was a significant main effect of Salience: $F(1, 39) = 19.21$, $p < .001$, $\eta_p^2 = .330$; children made number-based selections more frequently when the alternative choices were of low salience ($M = 2.29$, $SD = 1.83$) compared to high salience ($M = 1.24$, $SD = 0.99$). There was no significant main effect of Order nor a Salience × Order interaction, ps ≥ .42, η_p^2s ≤ .017, indicating that the main effect of Salience held for total match scores regardless of whether the AtN trials followed or preceded Set C novel trials.

Orientation-based matches. In the parallel set of analyses focused on orientation-based match selections, we conducted a 2 (Salience) × 2 (Order) mixed ANOVA on orientation-based best match scores. We replicated the significant main effect of Salience found in Chan and

Mazzocco (2017): $F(1, 39) = 36.63$, $p < .001$, $\eta_p^2 = .484$. Children made orientation-based selections more frequently when the alternative choices were of low salience ($M = 1.63$, $SD = 1.43$) compared to high salience ($M = 0.32$, $SD = 0.52$). In contrast with the results for number-based matches reported in the preceding section, there was a significant main effect of Order, $F(1, 39) = 8.14$, $p = .007$, $\eta_p^2 = .173$. Children who completed the AtN *after* completing Set C matched on orientation more frequently ($M = 1.29$, $SD = 0.77$) on the original AtN trials compared to children who completed the AtN *before* completing Set C ($M = 0.65$, $SD = 0.65$). However, there was a significant Salience × Order Interaction, $F(1, 39) = 7.84$, $p = .008$, $\eta_p^2 = .167$. The effect of Order differed across salience conditions; on low salience trials (combining orientation, pattern, and location), children who had first viewed Set C (which included number, orientation, and location trials alternating with color, shape, and pattern trials) made orientation-based best match selections more frequently ($M = 2.24$, $SD = 1.51$) compared to children who did not view Set C until *after* completing the original AtN trials ($M = 1.00$, $SD = 1.03$), $p = .004$. On high salience trials (orientation, shape, color), by contrast, there was no difference between children who viewed Set C first ($M = 0.33$, $SD = 0.48$) and those who viewed Set C last ($M = 0.30$, $SD = 0.57$), $p = .841$. These findings do *not* replicate our earlier findings (Chan & Mazzocco, 2017) per se, because Order effects were significant in the present study but not in our previous work. However, we did report an Order × Salience interaction in the original study (Chan & Mazzocco, 2017) and also in the present study, and that the interaction involved an effect of order on orientation- (but not number-) based responses in the low salience condition only, as a function of whether the AtN task began with trials that included combinations of options based on low salience features. Our general findings mirror the original AtN results.

The results for orientation-based matches partially held when the *total match* score was used instead of the *best match* score. Consistent with the best match score analyses, the significant main effect of Salience held, $F(1, 39) = 14.09$, $p < .001$, $\eta_p{}^2 = .265$, such that children made orientation-based selections more frequently when the alternative choices were of low salience ($M = 2.22$, $SD = 1.73$) compared to high salience ($M = 1.27$, $SD = 1.07$). The significant main effect of Order also held, $F(1, 39) = 14.55$, $p < .001$, $\eta_p{}^2 = .272$, and was in the same direction: Children who completed the AtN *after* completing Set C trials matched on orientation more frequently ($M = 2.33$, $SD = 1.00$) compared to children who completed the AtN *before* completing Set C trials ($M = 1.16$, $SD = 1.06$). However, the Salience × Order interaction was not significant, $F(1, 39) = 3.82$, $p = .058$, $\eta_p{}^2 = .089$.

Secondary analyses.

Novel combinations of features. A premise of the AtN task, and the motivation for the original study, is that *context* affects children's attention to numerosity. This notion is supported by the main effects of salience we reported earlier (Chan & Mazzocco, 2017), and our current replication of those findings. Perhaps, however, these salience effects are limited to the combinations of features appearing in the original AtN task used in both studies. By designing a novel combination of low salience features (i.e., Set C: orientation, location, number) to compare against the original set of high salience features (shape, color, number), we created an opportunity to test this notion further. Thus, we carried out a 2 (Salience: Set A1 vs. Set C1) × 2 (Order: Set C before or after Set B+A) mixed ANOVA on the number-based best match scores, and again found a significant main effect of Salience, $F(1, 39) = 23.34$, $p < .001$, $\eta_p{}^2 = .374$. Children chose number as the best match more frequently when alternative match choices were of low salience ($M = 1.44$, $SD = 1.29$) compared to high salience ($M = 0.37$, $SD = 0.80$). As with the original ANOVAs between Sets A and B, there was no significant main effect of Order, nor was there a Salience × Order interaction on the frequency of number-based matches, ps ≥ .18, ${\eta_p}^2$s ≤ .045.

Set C also allowed us to evaluate the relative frequency of all six feature-based responses in different combinations for descriptive purposes, which was not possible in the original AtN task. From this descriptive analysis only (Table 2), we did not observe shifts in number-based matches when pattern (B2) was replaced with orientation (C2). However, there were marked shifts in the relative frequency of orientation-based matches in Set C versus Set A. Children matched on orientation more frequently when the competing features included number instead of pattern. We also observed shifts in the relative frequency of shape- and color-based matches *to each other,* as a function of whether the third alternative match choice was number (Set A1), orientation (Set B1), or pattern (Set C1). The frequency of pattern-based matches was lower when the competing features were shape and color (Set C1) instead of location and orientation (Set A2) or location and number (Set B2).

AtN task and counting performance. Although all participants in the study had some number knowledge, the variation in their verbal counting, enumeration, and set production accuracy allowed us to explore potential relations between the frequency of number-based matches on the AtN and children's counting performance. Our first independent variable of interest was children's verbal counting level. We grouped children by whether their verbal counting range was below 20 (n = 16), between 20 and 40 (n = 15), or above 40 (n = 10), and examined whether their number-based best match or total match scores vary with verbal counting level. Based on two one-way ANOVAs, we found that children's best match or total match score did *not* differ as a function of verbal counting level, ps > .650, ${\eta_p}^2 < .018$. Our second independent variable of interest was children's enumeration accuracy, and we grouped children according to whether their highest set-counting accuracy was below six (*n* = 16), between six and nine (*n* = 12), or 10 (*n* = 13). Based on two one-way ANOVAs, children's best match and total match scores did *not* differ as a function of their enumeration accuracy level, *ps* > .744, ${\eta_p}^2 < .012$. Our last independent variable of interest was children's set production accuracy, and we grouped children according to whether their highest set production accuracy was below nine (*n* = 12), nine (*n* = 19), or 19 (*n* = 10). Based on two one-way ANOVAs, children's best match and total match scores did *not* differ as a function of their enumeration accuracy level, *ps* > .385, ${\eta_p}^2 < .049$.

Discussion

In Study 1, we replicated the main effects of salience on preschoolers' selection of both number-based and orientation-based matches, with a novel sample of preschoolers who were very similar to participants in our original study as reported by Chan and Mazzocco (2017). We demonstrated that performance on the AtN task was remarkably constant across the two studies. We also replicated the finding that number and orientation conform to similar outcomes and seem to be comparable in terms of their low salience status, at least in the context of the original AtN task.

Our novel combinations of features lent further support for the main effect of salience on the frequency of selecting number as a best match, showing once again that children select number-based matches more frequently when number appears in combination with other low salience features, compared to when it is combined with relatively high salience features. This was true also when the low salience set of trials (Set C) differed compared to the sets we used previously (Chan & Mazzocco, 2017). Moreover, this finding generalized to the feature *orientation*, which we included as a primary contrast to *number* (as another low salience feature), and in this way children's attention to number and orientation were similar. That is, children selected number and orientation to remarkably comparable degrees when each was pitted against color and shape (mean best match scores of 0.37 and 0.32, respectively, for number and orientation), or when pitted against pattern and location (mean best match scores of 1.49 and 1.63, respectively). However, when pitted against each other in novel Set C, number matches were less frequent than orientation matches (1.44 compared to 3.56, respectively), which was an unanticipated finding opposite that which we predicted. That orientation surpassed number in best match frequency in this novel combination suggests that number and orientation differ in their salience to each other, at least in the context of the AtN task, and further speaks to the remarkable infrequency with which children attend to number on this task.

Another outcome of including the novel trials was the ability to observe changes in the frequency with which children selected shape or color as a best match feature, as a function of what other (third) competing feature was among the choices present. Shape and color were selected with similar frequency when pitted against either number or orientation, in the present and original studies (Chan & Mazzocco, 2017). Yet when pitted against pattern instead of number or orientation (in Set C), the frequency of shape-based matches increased, and the frequency of color-based matches decreased, relative to Sets A and B (Table 2). Although this finding does not address the frequency with which children attend to numerosity specifically, it contributes to the general notion that competing features may influence to what children *do* attend. Accepting that selections for a best match do not reveal whether a child noticed number (or orientation), we relied on the total match scores to learn that the effects of salience were in the same direction when children revisited trials. The frequency of selecting number as a match

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

remained low across the sets, and children still matched on number less frequently when the competing features were of high vs. low salience, at least for Sets A and B.

These findings collectively reinforce the idea that number (i.e., numerosity), like the other features, is not a consistently salient feature across contexts—rather, its salience varies as a function of the salience of 'competing' visual features. The final step in our replication efforts was to replicate the main effect of salience we previously reported among adults (Chan & Mazzocco, 2017), which we pursued in Study 2.

Study 2

Method

Participants. Participants included 91 adults (31 men, 60 women; 64 white, 19 Asian or Pacific Islander, 3 black or African-American, 5 multiracial), all undergraduate (*n* = 86) or graduate students attending the University of Minnesota, who consented to participate in the study to receive either course credit or 10 US dollars. Their mean age was 21.2 (*SD* = 1.95) years, and they averaged 14.86 (*SD* = 1.16) years of formal education at the time of testing. One additional participant was removed from the current analyses for not completing one of the experimental tasks.

Materials.

Attention to Number (AtN) task. Adults completed the same paper-and-pencil version of the AtN task used in our original study (Chan & Mazzocco, 2017), as described in Study 1, consisting of the two original blocks of 16 trials (Sets A + B). There were several differences in the administration of the AtN task between Study 1 and Study 2, consistent with the differences reported for child and adult participant groups in the original study (Chan & Mazzocco, 2017). First, during the revisit trials, adult participants revisited all 32 trials (compared to only eight for child participants) and were asked to rank order their "next-best match" selections until they reported there were no additional matches (as opposed to reporting additional matches unranked). This afforded larger, and more informative *total match* scores (as in Study 1, these are computed as the total frequency of number-based matches during best match and revisit trials combined). In view of the significant effects we reported for the order in which Sets A and B were administered to adults (Chan & Mazzocco, 2017), Order was systematically varied (and counterbalanced) in Study 2 as a between-participants variable, with roughly half of the participants completing Set A trials first, and the other half completing Set B trials first. We prioritized capturing the more nuanced attention to number in Sets A and B in adults by asking them to rank order the matches, instead of adding Set C to Study 2. This also allowed us to directly replicate our original study with adults (Chan & Mazzocco, 2017).

Woodcock – Johnson III. Participants completed the 3-minute standardized Math Fluency subtest of Woodcock–Johnson III (WJ-III; Woodcock, McGrew, & Mather, 2001). This pencil-and-paper task requires participants to solve as many arithmetic problems on a worksheet as quickly as possible; problems appear in order of increasing difficulty. We recorded participants' accuracy (the number correct) and time to completion. The median test-retest reliability for this subtest, for adults, is .92 (Mather & Woodcock, 2002). Participants' fluency rate was calculated as the number correct per minute. One participant's score was omitted due to experimenter error, leaving 90 total scores.

ACT – Math. We requested access to participants' standardized scores through our university, and specifically collected data on the ACT – Math, which is the quantitative subset of the ACT standardized college entrance exam that has been used in the U.S. for over 60 years. The ACT – Math subtest is used to measure mathematical problem-solving on 60 multiple choice questions. In the midwestern regions of the U.S., the ACT is significantly more common compared to another college entrance exam, the SAT. Six participants had only SAT scores on file, so for these individuals we converted their SAT scores to ACT equivalents using the ACT's and SAT's published national percentile norms. ACT scores (or equivalents) were obtained for 72 of the 91 participants for whom we received permission and had access to official recorded scores from our university.

Procedure. All adult participants were tested individually in a dedicated University-based research office by one of three research assistants (including EB and TP-S), following a protocol approved by our institutional review board (IRB). Study 2 participants completed the AtN and WJ-III Math Fluency tasks in the context of a larger study (Brown, Mazzocco, Rinne, & Scanlon, 2016) in which we administered the AtN and WJ-III third and fourth among four tasks, respectively. The entire session took approximately one hour to complete. (We analyzed and reported on the first task in the earlier report by Brown and colleagues (2016) and excluded those tasks from the current study; conversely, the AtN task was excluded from the prior report.) Participants were naïve to the purpose of the experiment and were told it was a decision-making study in order to avoid biasing participants toward thinking about number.

Results

Preliminary descriptive analyses. We first examined the relative frequency with which each individual feature was selected as a best match across all AtN trials (Table 3), and found these frequencies to be comparable to those reported for the adults in Chan and Mazzocco's (2017) study, with adults choosing each competing feature at similar rates across the two studies (Table 3). Foils consistently received the lowest best match scores across trials (M = 1.35 out of 32 possible), and were selected at less than half the rate of the next lowest best match score

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

(orientation, *M* = 1.38 out of 16), suggesting that adults were not selecting match options randomly.

Primary analyses: replication.

Variables of interest. Consistent with Study 1 and our original AtN study (Chan & Mazzocco, 2017) that we aimed to replicate, our primary outcome variables were *attention to number* and (for comparison) *orientation*, as measured by the frequency with which adult participants made number-based (and orientation-based) matches on *best match* trials (i.e., their best match score) or on any trials including the initial plus revisit trials (i.e., their total match score). Our primary independent variable of interest was the two levels of salience manipulated across trials, corresponding to whether the features that competed with number-based (or orientation-based) matches were classified as *high* or *low* salience. Our second independent variable of interest was the order in which Sets A and B were administered, which we measured in an effort to replicate the Salience × Order interaction we previously reported for adults (but not children; Chan & Mazzocco, 2017). As with Study 1, we used mixed analyses of variance (ANOVAs) for our primary analyses and Bonferroni-corrected pairwise comparisons for significant interactions. We conducted parallel mixed ANOVAs for number-based and orientation-based selections, each using best match and total match scores, separately. All analyses were performed in R (R Core Team, 2019) using Tidyverse (Wickham et al., 2019).

Adults' attention to number and orientation.

Number-based matches. Using a 2 (Salience of competing features, within-participants) × 2 (Order of Sets A and B, between-participants) mixed ANOVA on number-based best match scores (Figure 2a), we replicated the significant main effect of Salience, $F(1, 89) = 65.52$, $p < .001$, $\eta_p{}^2 = .424$. Adults made number-based best match selections more frequently when the alternative choices were of low salience (*M* = 2.29 out of 8 possible, *SD* = 2.46) compared to high salience (*M* = 0.54, *SD* = 1.10). Unlike in our previous findings (Chan & Mazzocco, 2017), the main effect of Order *was* statistically significant, $F(1, 89) = 7.33$, $p = .008$, $\eta_p{}^2 = .076$, with adults making more number-based best matches overall when low salience trials were viewed first (*M* = 1.86, *SD* = 2.37) instead of second (*M* = 1.01, *SD* = 1.71). We replicated the statistically significant Salience × Order interaction we previously reported (Chan & Mazzocco, 2017), $F(1, 89) = 10.77$, $p = .001$, $\eta_p{}^2 = .108$.

We ran a post-hoc analysis of this interaction, based on Welch's *t*-tests with Bonferroni corrections, and found that for trials with low-salience competing features (pattern and location), participants who saw number with low-salience features in their first block of trials chose number as a best match significantly more frequently (*M* = 3.12, *SD* = 2.59) compared to participants who saw number with high-salience features in their first block of trials (*M* = 1.54, *SD* = 2.08), $t(80.5) = 3.17$, $p = .004$. By contrast, for trials with high-salience competing features, participants who saw those high-salience trials first did *not* show different rates of choosing number as a best match (*M* = 0.48, *SD* = 1.01) compared to participants who had seen low-salience trials first (*M* = 0.60, *SD* = 1.20), $t(82.6) = 0.54$, $p = 1.00$. This replicates the values and differences we observed in our original study (Chan & Mazzocco, 2017).

The results for number-based matches partially held when the *total match* score was used as the dependent variable (Figure 2b). Consistent with the best match score analyses, the significant main effect of Salience held, $F(1, 89) = 10.16$, $p = .002$, $\eta_p{}^2 = .102$, with adults selecting number-based matches more frequently when alternative choices were of low salience (*M* = 4.16, *SD* = 3.00) compared to high salience (*M* = 3.38, *SD* = 3.29). The significant main effect of Order also held, $F(1, 89) = 5.08$, $p = .027$, $\eta_p{}^2 = .054$, with adults making number-

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

based matches more frequently when low salience trials came first (M = 4.49, SD = 3.17) compared to when high salience trials came first (M = 3.14, SD = 3.04). However, unlike this sample's best match score or the total match scores we reported earlier (Chan & Mazzocco, 2017), there was no significant Salience × Order interaction on total match scores, p = .567, $\eta_p{}^2$ = .004.

Orientation-based matches. In the parallel set of analyses focused on orientation-based match selections, we conducted a 2 (Salience) × 2 (Order) mixed ANOVA on the frequency of orientation as a best match (Table 4). Here we replicated the significant main effect of Salience we previously found (Chan & Mazzocco, 2017), $F(1, 89) = 28.19$, $p < .001$, $\eta_p{}^2 = .241$, with participants making orientation-based best matches significantly more frequently when competing features were of low salience (M = 1.04 out of 8, SD = 1.11) compared to high salience (M = 0.34, SD = 0.65). Also consistent with the original study, the main effect of Order was not statistically significant, $F(1, 89) = 0.06$, $p = .813$, $\eta_p{}^2 < .001$. However, unlike in our earlier study (Chan & Mazzocco, 2017) there was no significant Salience × Order interaction, $F(1, 89) = 1.24$, $p = .269$, $\eta_p{}^2 = .014$.

The results held when the *total match* score replaced the *best match* score as the dependent variable (Table 4). There was a significant main effect of Salience, $F(1, 89)$ = 4.79, $p = .031$, $\eta_p{}^2 = .051$, with adults selecting orientation-based matches more frequently when competing features were of low salience (M = 1.45, SD = 1.39) compared to high salience (M = 1.07, SD = 1.50). The main effect of Order was not statistically significant, $F(1, 89) = 0.01$, p = .912, $\eta_p{}^2 < .001$; nor was the Salience × Order interaction, $F(1, 89) = 3.01$, $p = .086$, $\eta_p{}^2 = .033$.

Secondary analyses: individual differences in Attention to Number. In our original AtN study (Chan & Mazzocco, 2017), we reported wide variation in attention to numerosity

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

across the adults, and provided descriptive information on the breadth of this variation without exploring its association with mathematics achievement. In the present study, we conducted an exploratory analysis of adults' bivariate frequencies across low- and high-salience trials (Figure 3) as a way to better understand the nature of their performance differences. We chose total match scores for this analysis due to the floor effects on best match scores (especially against high-salience competing features), and used statistical significance tests to categorize participants into three groups based on whether they chose number-based matches across best match plus revisit trials (a) rarely, (b) frequently but only when the competing features were of low- compared to high-salience, and (c) frequently regardless of the salience of competing features.

To determine chance probability of matching based on number, we accounted for the dependencies involved (e.g., choosing number as a second choice depends on not choosing number as the best match), and the fact that there is an option to choose "no more matches" on each revisit trial:

$$P(N) = P(N_1) + P(N_2, \neg N_1) + P(N_3, \neg N_1, \neg (N_2 \vee \emptyset_2))$$

where N_1 represents a choice of the number match first, $\emptyset_2$ represents a choice of "no more matches" in the second choice-point, etc. Applying the chain rule, this is computed by:

$$P(N) = P(N_1) + P(N_2 \mid \neg N_1)P(\neg N_1) + P(N_3 \mid \neg N_1, \neg (N_2 \vee \emptyset_2))P(\neg (N_2 \vee \emptyset_2) \mid \neg N_1)P(\neg N_1).$$

Since there are four match options for best match, four options for second choice (three remaining options + "no more matches"), and three options for third (two remaining + "no more matches"), chance is calculated as:

$$P(N) = (.25) + (.25 * .75) + (.\overline{3} * .5 * .75) = .5625.$$

Thus, the probability of ever choosing a number-based match by chance is .5625, or 9/16.

Using the binomial distribution with $H_0: P(N) = .5625$, total match scores of 0 to 4 out of 16 trials total are significantly *below* chance ($p < .05$) and total match scores of 13 to 16 out of 16 are significantly *above* chance, all *p*s two-tailed. Thus, participants with total match scores of 4 or lower were classified as choosing number rarely, and those of 13 or higher as choosing number frequently and independent of context (see Figure 4).

In order to classify participants as being context-sensitive, we used Fisher's exact test to determine what frequencies of number-based total match scores during low salience trials (out of 8) are significantly higher compared to frequencies of number-based total match scores during high salience trials (also out of 8). We found that participants must have at least 5 additional total match scores *from low salience trials* compared to high salience trials (e.g., 5/8 on low salience trials and 0/8 on high salience trials, or 6/8 on low salience trials and 1/8 on high salience trials) to meet this classification criterion. Thus, all participants with those bivariate frequencies are classified as choosing number more frequently in low salience conditions.

Using these classifications (Figure 4), 38 participants were categorized as choosing number rarely (significantly below chance), seven as contextually-sensitive (significantly more in low-salience than high-salience), and 27 as frequently regardless of context (significantly above chance), and 19 participants demonstrated inconsistent patterns not otherwise classified.

Although this classification scheme was exploratory, we also investigated whether classifications were influenced by which AtN trials adults completed first. A simulated power analysis using the 'simpr' R package (Brown & Bye, 2020) found the minimum odds ratio to detect at 80% power is 4.33 for our sample sizes. Noting that we had the statistical power to detect only such a large effect, we next used Fisher's exact test to compare the proportion of participants who saw the low-salience trials first (Set B2) and were categorized as choosing number either rarely or frequently (contextually-independent), compared to those who saw the high-salience trials first (Set A1). We found no statistically significant difference, OR = 2.31, p = .13, 95% CI = [0.76, 7.31], although the *low-salience first* condition had a descriptively higher proportion of participants choosing number frequently (17 out of 33, 52%), relative to the proportion for adults who completed the high-salience trials first (10 out of 32, 31%).

Correlations with math achievement scores. We used ANOVA to also examine whether group classification (rare, context-sensitive, frequent, or inconsistent) was related to participants' scores on the Woodcock Johnson – III Math Fluency (n = 90, M = 47.1, SD = 10.5) or ACT – Math score (n = 72, M = 27.5, SD = 4.3). There was no main effect of group classification on math achievement from either ANOVA, Fs < 1.6, ps > .21.

We next examined the relation between participants' total number match scores and the two mathematics performance measures. Total match scores were significantly correlated with ACT – Math scores, $r(70)$ = .25, p = .033, 95% CI: [.02, .46], but not with Woodcock Johnson – III Math Fluency, $r(88)$ = -.03, p = .79, 95% CI: [-.23, .18]. In order to also compare the salience effect to these measures, we ran a multiple regression on each outcome to examine whether the total match score, context-specific total match score (total match score for low salience trials – high salience trials), or their interaction predicted either performance measure. We included the interaction between total match score and context-specific total match score in order to address whether the association between total match score and mathematics performance depended, at least partially, on participants' sensitivity to salience. In the full model, total match score, context-specific match score, and their interaction were not significant predictors of either performance measure, ts < 1.56, ps > .12.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

Finally, we weighted each total match score by the order in which a participant matched on number—e.g., a best match was coded as 4, a second-round match as 3, third-round as 2, fourth-round as 1, and no number match as 0. We then performed a parallel set of analyses comparing these 'index-weighted' match scores to each mathematics performance measure. We again found a significant correlation with ACT – Math scores, $r(70) = .24$, $p = .038$, 95% CI: [.01, .45], but not with Woodcock Johnson – III Math Fluency, $r(88) = .01$, $p = .94$, 95% CI: [-.20, .21]. The multiple regressions using both total index-weighted match score, context-specific index-weighted match score, and their interaction again revealed no significant predictors, ts < 1.92, ps > .06.

Discussion

In Study 2, we replicated the main effect of salience on the frequency of adults' number-based (and orientation-based) picture match responses reported in our original study (Chan & Mazzocco, 2017), as follows: Adults made number-based (and orientation-based) selections more frequently when competing match options were based on features of relatively low compared to high salience. In doing so, we demonstrated that AtN task performance was remarkably constant across the two studies of adults. As we reported for children in Study 1, the infrequency with which adults selected foils as matches counters notions that participants were choosing randomly. Finally, we replicated the main effect of Order reported in our original study (Chan & Mazzocco, 2017), showing that adults who began the AtN with a block of trials in which number matches competed with low salience features were more likely to select number-based matches overall, compared to adults who began the AtN with a block of trials in which number matches competed with high salience features. We also replicated the Order by Salience moderating interaction, suggesting that earlier exposure to low-salience competing features increases the amount of number-based best matches on trials with low-salience features to a greater extent than for trials with high-salience features. This sort of priming for numerosity-based matches may be another demonstration of the effects of context on attention to numerosity.

We were able to explore individual differences in attention to numerosity more thoroughly in Study 2 compared to Study 1, owing to the wider range of scores among adults compared to children and, potentially, the larger sample size in Study 2. Adults showed considerable individual differences as to whether they never, sometimes, or regularly matched based on numerosity, but also on whether they were subject to effect of context across conditions. The more straightforward finding concerned a significant but small correlation between AtN scores and the broad math achievement score (ACT-Math), which did not emerge for the specific skill of math fluency. Still, this correlation was small (effect size of .056), and may reflect the association between collective mathematics experiences in school and attention to numerosity in the AtN task or an underlying tendency to focus on numerosity. Neither explanation can be ruled out, nor supported, based on this small effect. Salience (low versus high) was a within-participants variable, order was a between-participants variable, and both were subject to main effects and a significant interaction. Therefore, since overall attention to numerosity was influenced by which block of trials adults completed first, it was more challenging to definitively classify individuals along a dimension representing their relative "context sensitivity" with respect to attending to numerosity. Nevertheless, the main effects of salience and order both speak to the effects of context on whether adults attended to numerosity on the AtN task.

General Discussion

The AtN task was designed to help determine *to which features children attend* among a finite set and how co-occurring features influence individuals' attention to numerosity specifically and to low salience features generally (i.e., numerosity, orientation; Chan & Mazzocco, 2017). In Studies 1 and 2, we demonstrated that AtN task performance was remarkably constant across children and adults, and across the present study and earlier findings (Chan & Mazzocco, 2017). This replication is important for demonstrating reliability of the AtN task; it also supports continued AtN task research as a means to better understand SFON and to uncover what traditional SFON tasks may be capturing in general and in terms of their predictive relation with later mathematics achievement.

This replicated main effect of salience supports the notion that context affects children's and adults' attention to numerosity relative to co-occurring visual features. Replicating the order effect among adults suggests that the influence of context may carry over to subsequent trials (at least within a task). Although order effects were not evaluated in the same manner with child participants (because of the fixed order of trial blocks in the original AtN task they completed), the order in which the novel set of trials (number, orientation, location) was presented had no effect on children's attention to numerosity, but did affect their attention to orientation. If we interpret the order effect to reflect priming susceptibility, adults appear more susceptible to such priming for numerosity compared to children, which may be related to a more mature awareness of numerosity from years of schooling. However, children's overall frequency of attention to numerosity was so low that it was unlikely any priming could occur.

Number and Orientation

The fact that children selected orientation matches more frequently when the novel trials preceded the original AtN task sheds light on the general principle that a feature of very low salience can be manipulated to appear relatively more salient even among young children. Among these children, orientation-based matches were selected in Set C trials *more than twice as often* than in Set A trials (when orientation was combined with low salience features), and *ten times more often* than in Set B trials (when orientation was combined with color and shape, the high salience competing features; Table 2). One explanation for why orientation, but not number, was susceptible to this priming may be that orientation-based matches outnumbered number-based matches in the novel Set C2 trials (when both are present in the same trial). That is, although number and orientation responses were similar to each other (in frequency and rank) across conditions of the AtN task, children were more likely to attend to orientation (compared to number) when these two features occurred together in Set C2 trials, making it conform to a relatively *more* salient feature at least in this particular context.

Among preschool children, numerosity, as a feature, does not stand out among most other features we examined. Adults are more likely to attend to it and, therefore, are more susceptible to its potential priming effects on a matching task. A potential application of such priming may be relevant to the development of early instructional materials (such as counting books or educational games) designed to support children's greater attention to number, and is consistent with prior research demonstrating the instructional value of plain compared to colorful or highly distracting learning environments (Hanley, Khairat, Taylor, Wilson, Cole-Fletcher, & Riby, 2017) or math manipulatives (McNeil, Uttal, Jarvin, & Sternberg, 2009). Our participans' focus on color supports prior evidence suggesting that mathematical learning materials are more effective when they are monochromatic, and extends this idea to early learning materials designed to support children's learning of numerosity, number words, and counting.

Although comparing number- and orientation-based matches was not a primary focus of the study, it is worth noting that we replicated most of our earlier findings (Chan & Mazzocco, 2017) absent the novel combinations, in new samples of children and adults. On the one hand, this replication suggests that these two features are appropriate substitutes for each other as features of relatively low salience, at least in contexts such as the AtN task. On the other hand, our replicated findings emphasize important *differences* between these two features. Orientation (but not number) was subject to an Order effect in our current study (and in our earlier work reported by Chan and Mazzocco, 2017) for both children and adults, suggesting greater susceptibility to priming effects for orientation compared to number. Number- and orientation-based matches occurred with remarkably similar frequency among children (but not among adults) in both our current study and our original study. Yet orientation-based matches occurred more frequently in our novel combination of trials used in Study 1, at least based on descriptive data summarized in Table 2. This supports the notion of orientation being susceptible to effects of task contexts – it met criteria for relatively low to relatively high salience, depending on the features with which it was paired.

Number was not so vulnerable. As seen in Table 2, children made fewer number-based matches than almost all non-foil alternatives across trials. This raises many questions about the salience of number in children's everyday play and instructional environments, and what features of traditional SFON tasks successfully draw children's attention to numerosity. This unexpected finding also has implications for the AtN task itself, because the shifts across Sets in relative frequency with which certain features are selected demonstrate the effect of visual task contexts on selecting features in general, consistent with our premise that task context matters. Still, these shifts raise questions about the conditions under which the main effects of salience on attention to numerosity will persist. What remains remarkable is the extent to which number continues to be selected so infrequently, relative to competing features, across studies, ages, and combinations of features. This suggests the importance of finding ways to draw children's attention to numerosity, if SFON as a construct is predictive of mathematics achievement.

What the AtN task may reveal about SFON

The AtN task is clearly not a traditional SFON task. It was designed to be a minimalistic, streamlined task wherein numerosity is a *potential* feature to which individuals may attend. Like several original SFON tasks (e.g., Hannula & Lehtinen, 2005), the AtN involves no explicit numerical prompting. The similarities between the AtN and traditional SFON tasks may end there, however, and important differences are intended.

First, although traditional SFON tasks obviously occur within the context of additional features naturally present in any environment (e.g., the color and shape of the materials), in the AtN task, that variation in competing features is systematically controlled in order to understand the potential sources of implicit prompts. Both traditional SFON tasks and the AtN task deliberately exclude verbal cues related to numerosity, but the AtN task is designed to also avoid nonverbal cues related to any one feature, whereas many traditional SFON tasks include such cues, such as sounds associated with items dropping into containers or stamps placed on papers. Compared to the frequency of numerosity, in the AtN task the other competing features occur *with equal frequency*, and are visible in the same locations (across trials), and for similar amounts of time, since the AtN is an experimental task designed to control for these factors.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

Second, the AtN task involves no numerically-relevant *actions* on the part of the experimenter or participants, whereas many SFON tasks may involve, for instance, feeding a toy animal, posting letters in a mailbox, or stamping pictures. Although not all SFON tasks require such actions, and the action-based condition we created (Chan & Mazzocco, 2017) did not influence attention to numerosity on the AtN task, the role of actions themselves on children's attention to numerosity remains an empirical question, as do other implicit features of SFON tasks. These implicit prompts may play an important role in instructional or play materials that may inadvertently raise or diminish the likelihood that a child notices numerosity. For instance, our findings suggest that in games that involve matching cards, blocks, puzzle pieces, or other materials, limiting the match criteria to numerical and other low salient features may encourage more attention to numerosity compared to materials that allow for matching based on color or shape. Findings from traditional SFON tasks suggest that the presence of multiple cues – such as correlated visual and auditory signals – may elicit attention to numerosity. So whereas the AtN task allows us to measure whether children can attend to numerosity *when it is an option among many other features*, traditional SFON tasks may measure whether children notice numerosity in situations that increases its salience—such as when a child observes letters being mailed and hears each time a letter drops into the mailbox.

Finally, scoring differences between the AtN and traditional SFON tasks include the exclusive focus on attention to *specific* set size in the former and *any* reference to numerosity in some cases of the latter. These differences afford an opportunity to study potential sources of individual differences across types of numerosity-based responses. A challenge to the self-initiated nature of the SFON construct is our finding that task *contexts* accounted for much variation in numerosity-based responses. Perhaps what varies across children engaged in the AtN

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

or other SFON tasks is the contexts with which children are familiar and how children's familiarity with the materials used in the tasks may prime some children to pay attention to numerosity (e.g., if they understand that certain objects are more likely to be counted than others).

Individual Differences

The descriptive data on individual differences from Study 1 and 2 mirror the descriptive report from our original study (Chan & Mazzocco, 2017). In the present study, we extended those findings to provide interesting new insight into adults' attention to numerosity and a potential avenue for future research. On average, adults choose number-based best matches relatively infrequently ($M = 2.82$ out of 16 trials), and only use some opportunities to identify a number-based match during the initial and revisit trials combined ($M = 7.55$ out of 16). Across participants, however, considerable variability reflected by large standard deviations (3.09, 5.85, for best and total match scores, respectively) raises questions about what differentiates adults who *rarely* choose number-based matches (i.e., those depicted on the upper left corner of Figure 4) from those who choose number-based matches frequently, either regardless of context (lower right corner) or only when alternative match options are of low salience (upper right).

Even when provided the opportunity to revisit AtN trials, adults often do *not* match on number, especially in the presence of higher salience features. Yet in our exploratory classification of participants according to statistical significance regions (Figure 4), we see that almost a third of participants are significantly above chance in number-based total match scores (27 of 91), over a third are significantly below chance ($n = 38$), and fewer are significantly more likely to match on number in the presence of low salience features compared to high ($n = 7$). That almost 80% (72 of 91) of participants fall into these categories suggests that individual differences are worth examining more directly in future research, which should measure the reliability of these classifications in larger samples, with more trials, and across a greater range of contexts and set orders.

We did not classify children into these categories because of our small sample and their small number of revisit trials in Study 1. We did, however, observe descriptive similarities to individual differences we originally reported (Chan & Mazzocco, 2017)—that many children never matched on numerosity, some (very few) children frequently matched on numerosity on only four or fewer of the eight high salience trials (Set A1), or on six or more of the eight low salience trials (Sets C2 and B2, respectively). The primary difference between children's and adults' variation is the relatively high rate of skewing to *never* matching on numerosity among children, and the equivalent rate of rarely and frequently matching on numerosity in adults. This comparison is tentative, however, because adults had four times as many opportunities to revisit trials than did children.

Adults' more frequent attention to numerosity in general, and thus the wider range of their number-based match scores, may explain why we begin to see correlations between AtN scores and math in adults but not children. Still, these correlations—when they occurred—were small, and would need replication to warrant further interpretation.

Conclusions and Future Directions

As with the original AtN study, our study supports the notion that task context affects attention to numerosity, a notion that may have implications for designing early childhood mathematics or play materials. The contextual effects we describe in this study may initially appear to challenge a defining feature of SFON—its spontaneity—which in turn raises the question of whether SFON reflects children's *tendencies* or their learning environment or history. We do not view these alternative explanations as mutually exclusive, because the factors underlying children's SFON performance may vary throughout development, or at least between early childhood and adulthood. That is, perhaps the *relative* contribution of context and natural tendencies is important for understanding *when* children attend to numerosity, *at what point* in their development they benefit from implicit cues (i.e., "SFON baits" discussed in previous studies, e.g., McMullen et al., 2019), and when their performance on the AtN or on traditional SFON tasks represents a response pattern foundational for learning mathematics. These questions may serve as important guides to future studies of children's attention to numerosity. Performance on the AtN task may be viewed as a good reflection of SFON *if* we consider the task an extremely streamlined measure of attention to numerosity; of course others may interpret the AtN as a task that does *not* measure SFON because it lacks a minimum threshold needed to *implicitly* cue for attention to numerosity despite the absence of numerical verbal prompts, as with traditional SFON tasks. We argue that the AtN task objectively measures attention to numerosity but not its spontaneity, because we deliberately but implicitly prime for attention to numerosity by pitting number against less salient features. Perhaps the continuum of SFON tasks is as follows: high salience trials of the AtN task may be the most basic measure of SFON, because these trials involve no cues or primes to attend to numerosity; low salience trials of the AtN task are a measure of *primed* focus on number; and traditional SFON tasks may actually measure children's receptiveness to nonverbal cues to focus on number. These notions and others proffered in the special issue are subject to empirical pursuit. In the meantime, our results offer explicit support for the theoretical motivation underlying the design of the AtN task, that aspects of materials or other task characteristics provide implicit prompts for attention to number even in the absence of explicit prompts for such attention.

References

Baroody, A. J., Li, X., & Lai, M. (2008). Toddlers' spontaneous attention to number. *Mathematical Thinking and Learning, 10*, 240–270.

Batchelor, S., Inglis, M., & Gilmore, C. (2015). Spontaneous focusing on numerosity and the arithmetic advantage. *Learning and Instruction*, *40*, 79–88.

Brown, E. C., & Bye, J. K. (2020). simpr: Flexible Tidyverse-Friendly Simulation Interface. R package version 0.1-0. https://github.com/statisfactions/simpr

Brown, E. C., Mazzocco, M. M. M., Rinne, L. F., & Scanlon, N. S. (2016). Uncanny sums and products may prompt "wise choices": Semantic misalignment and numerical judgments. *Journal of Numerical Cognition*, *2*(2), 116–139. https://doi.org/10.5964/jnc.v2i2.21

Chan, J. Y. C., & Mazzocco, M. M. (2017). Competing features influence children's attention to number. *Journal of Experimental Child Psychology*, *156*, 62-81.

Chan, J. Y. C., Praus-Singh, T. L., & Mazzocco, M. M. (2020). Parents' and young children's attention to mathematical features varies across play materials. *Early Childhood Research Quarterly*, *50*, 65-77. https://doi.org/10.1016/j.ecresq.2019.03.002

Clark, A., Henderson, P., & Gifford, S. (2020). *Improving Mathematics in the Early Years and Key Stage 1.* Retrieved from Education Endowment Foundation website: https://educationendowmentfoundation.org.uk/public/files/Publications/Maths/EEF_Maths_EYKS1_Guidance_Report.pdf

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLOS ONE*, *8*(1), e54651.

Hanley, M., Khairat, M., Taylor, K., Wilson, R., Cole-Fletcher, R., & Riby, D. M. (2017). Classroom displays—Attraction or distraction? Evidence of impact on attention and learning from children with and without autism. *Developmental Psychology, 53*(7), 1265–1275. https://doi.org/10.1037/dev0000271

Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*, 237–256. https://doi.org/10.1016/j.learninstruc.2005.04.005

Hannula-Sormunen, M. M., Lehtinen, E., & Räsänen, P. (2015). Preschool children's spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Mathematical Thinking and Learning*, *17*, 155–177. https://doi.org/10.1080/10986065.2015.1016814

Hannula, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, *107*, 394–406. https://doi.org/10.1016/j.jecp.2010.06.004

Hannula, M. M., Mattinen, A., & Lehtinen, E. (2005). Does social interaction influence 3-year-old children's tendency to focus on numerosity? A quasi-experimental study in day care. In L. Verschaffel, E. De Corte, G. Kanselaar, & M. Valcke (Eds.), *Powerful Environments for Promoting Deep Conceptual and Strategic Learning* (pp. 63–80). Leuven, Belgium: Leuven University Press.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850-867.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299-321.

Mather, N., & Woodcock, R. W. (2001). *Examiner's manual: Woodcock-Johnson III tests of achievement*. Itasca, IL, USA: Riverside Publishing.

Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, *20*(3), 142-155.

McCloskey, M. (2007). Quantitative literacy and developmental dyscalculias. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 415-429). Baltimore, MD, US: Paul H Brookes Publishing.

McMullen J., Chan J.YC., Mazzocco M.M.M., Hannula-Sormunen M.M. (2019). Spontaneous Mathematical Focusing Tendencies in Mathematical Development and Education. In A. Norton & M. W. Alibali (Eds.), *Constructing Number: Merging Perspectives from Psychology and Mathematics Education* (69-86). Springer, Cham.

McMullen, J. A., Hannula-Sormunen, M. M., & Lehtinen, E. (2013). Young children's recognition of quantitative relations in mathematically unspecified settings. *The Journal of Mathematical Behavior*, *32*(3), 450-460.

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learning and Instruction*, *19*(2), 171–184. https://doi.org/10.1016/j.learninstruc.2008.03.005

National Governors Association Szekely, A. Unlocking Young Children's Potential: Governors' Role in Strengthening Early Mathematics Learning. Washington, D.C.: National Governors Association Center for Best Practices, October 28, 2014.

R Core Team (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Verschaffel, L., Rathé, S., Wijns, N., Degrande, T., Van Dooren, W., De Smedt, B., & Torbeyns, J. (2018, May). Young children's early mathematical competencies: The role of mathematical focusing tendencies. In I. Erfjord (Ed.), *Mathematics education in the early years. Results from the POEM4 Conference, 2018.* Springer.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

Table 1

Means and Standard Deviations (in Parentheses) for Frequency of Children's Best-Match Response by Feature [a] *on the Original AtN Trials*

	Shape	Color	Number	Orientation	Pattern	Location	Foil
Current study	7.71 (5.27)	7.24 (5.42)	1.85 (2.21)	1.95 (1.55)	8.29 (4.19)	2.78 (2.41)	2.17 (1.55)
Chan & Mazzocco (2017)	6.88 (5.76)	7.68 (6.08)	1.28 (0.64)	2.56 (1.44)	6.72 (3.04)	3.68 (1.60)	3.20 (2.56)

Note. The data in row 2 are drawn from "Competing Features Influence Children's Attention to Number," by J. Y.-C. Chan and M. M. M. Mazzocco, 2017, *Journal of Experimental Child Psychology, 156,* p. 69. Copyright 2017 by Elsevier.

[a] The maximum number of trials per feature is 16, and for Foils it is 32.

Table 2

Means and Standard Deviations (in Parentheses) for the Frequency of Preschoolers' Best Match Responses, by Feature, Across Eight AtN Trials, Per Feature Combination from the Original AtN Task (Sets A and B) and Novel Combinations (Set C)

Subset	Shape	Color	Number	Orientation	Pattern	Location	Foil
A1: SCNF	3.95 (3.07)	3.49 (3.02)	0.37 (0.80)				0.20 (0.51)
A2: PLOF				1.63 (1.43)	3.63 (2.23)	1.76 (1.67)	0.98 (0.82)
B1: SCOF	3.76 (2.48)	3.76 (2.69)		0.32 (0.52)			0.17 (0.50)
B2: PLNF			1.49 (1.61)		4.66 (2.29)	1.02 (1.15)	0.83 (0.89)
C1: SCPF	4.81 (3.20)	1.24 (2.13)			1.34 (2.07)		0.61 (1.36)
C2: NOLF			1.44 (1.29)	3.56 (1.40)		1.73 (1.20)	1.27 (1.25)

Note. N = 41. Features are indicated as follows: S=Shape; C=Color; N=Number; P=Pattern; L=Location; O=Orientation. F=Foil. Each subset of trials is comprised of 8 trials, each with a combination of three features and a foil. Blank cells indicate values for which no data are applicable.

Table 3

Means and Standard Deviations (in Parentheses) for Frequency of Adults' Best-Match Response by Feature[a] on the Original AtN Trail

Sample	Shape	Color	Number	Orientation	Pattern	Location	Foil
Current	11.91 (3.95)	3.07 (3.34)	2.82 (3.09)	1.38 (1.29)	9.51 (4.58)	1.96 (2.29)	1.35 (1.26)
Chan & Mazzocco (2017)	10.72 (4.32)	3.68 (3.36)	3.36 (3.20)	1.76 (2.40)	8.96 (4.48)	2.56 (2.88)	0.96 (1.28)

Note. The data in row 2 are drawn from "Competing Features Influence Children's Attention to Number," by J. Y.-C. Chan and M. M. M. Mazzocco, 2017, *Journal of Experimental Child Psychology, 156,* p. 69. Copyright 2017 by Elsevier.

[a] The maximum number of trials per feature is 16, and for Foils it is 32.

Table 4

Descriptive Statistics for Adults' Orientation-Based Responses Across Eight Trials

	Best Match		Total Score		*n*
Condition	*M*	*SD*	*M*	*SD*	
High Salience First					
High Salience	0.28	0.59	0.88	1.42	43
Low Salience	1.14	1.26	1.60	1.50	43
Low Salience First					
High Salience	0.40	0.71	1.23	1.56	48
Low Salience	0.96	0.94	1.31	1.29	48

Note. Competing features in high salience trials were shape and color. Competing features in low salience trials were pattern and location.

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

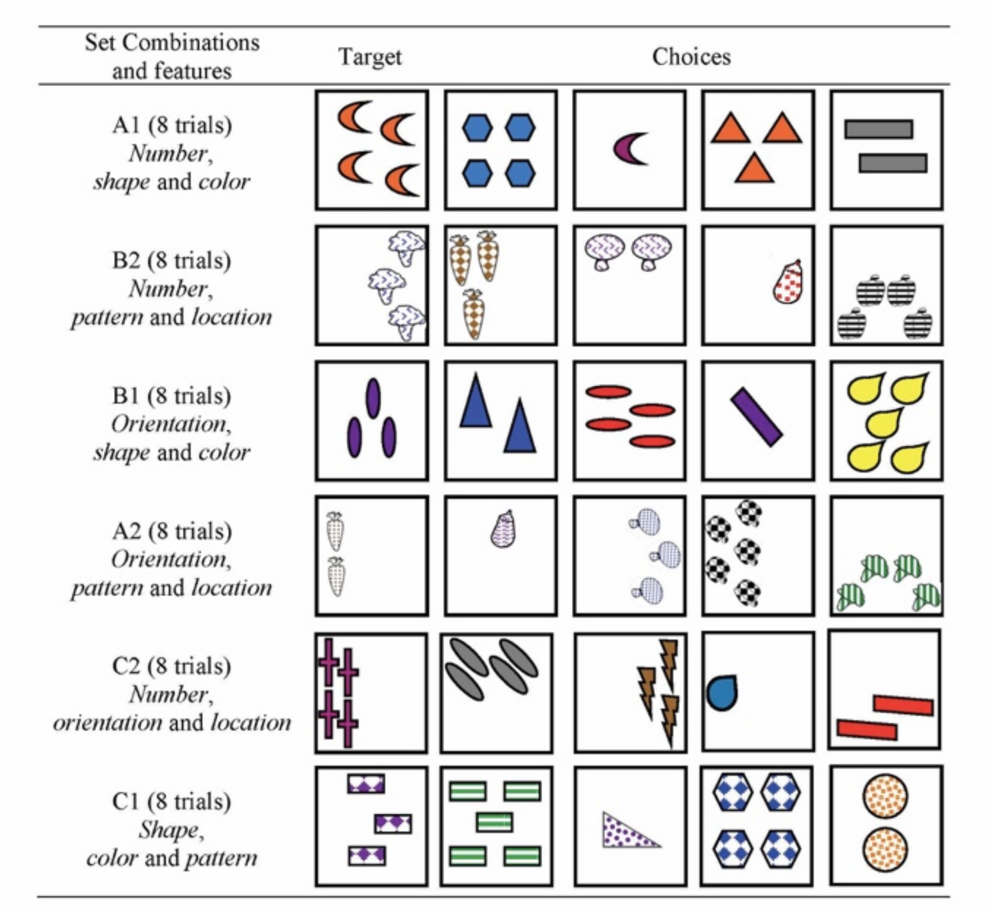

Figure 1. Exemplars of each of the six stimulus combinations used in the present study. Sets A and B are from the original Attention to Number task. Set C is unique to this study. Each combination includes one target and four response choices, including a foil. Although features appear in the same relative location in this figure, with foils always appearing last, the location of the features/foils is counterbalanced across pages of the test booklet. Adapted from "Competing Features Influence Children's Attention to Number," by J. Y.-C. Chan and M. M. M. Mazzocco, 2017, *Journal of Experimental Child Psychology, 156,* p. 66. Copyright 2017 by Elsevier.

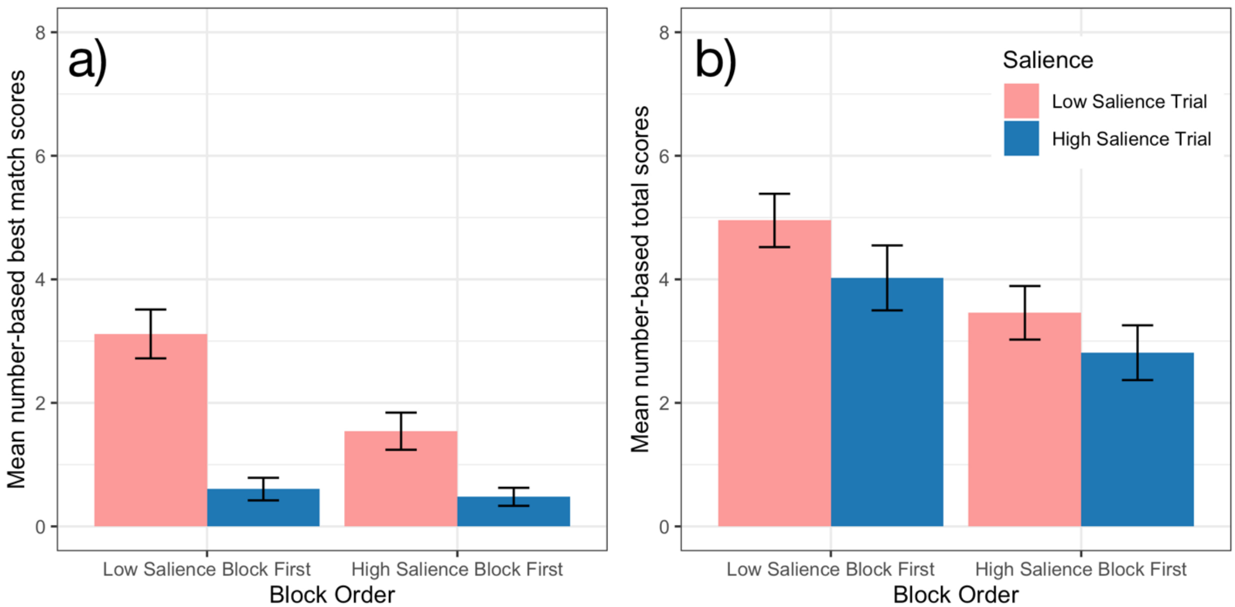

Figure 2. The mean (and standard error) of number-based best match scores (a) and total scores (b) for adults in Study 2, as a function of whether trials have low (pattern and location) or high (color and shape) salience competing features and the order in which participants saw each block (low or high salience trials first).

This is an Accepted Manuscript of an article published by Taylor & Francis in Mathematical Thinking and Learning on 18 Sep 2020, available at: https://doi.org/10.1080/10986065.2020.1818467

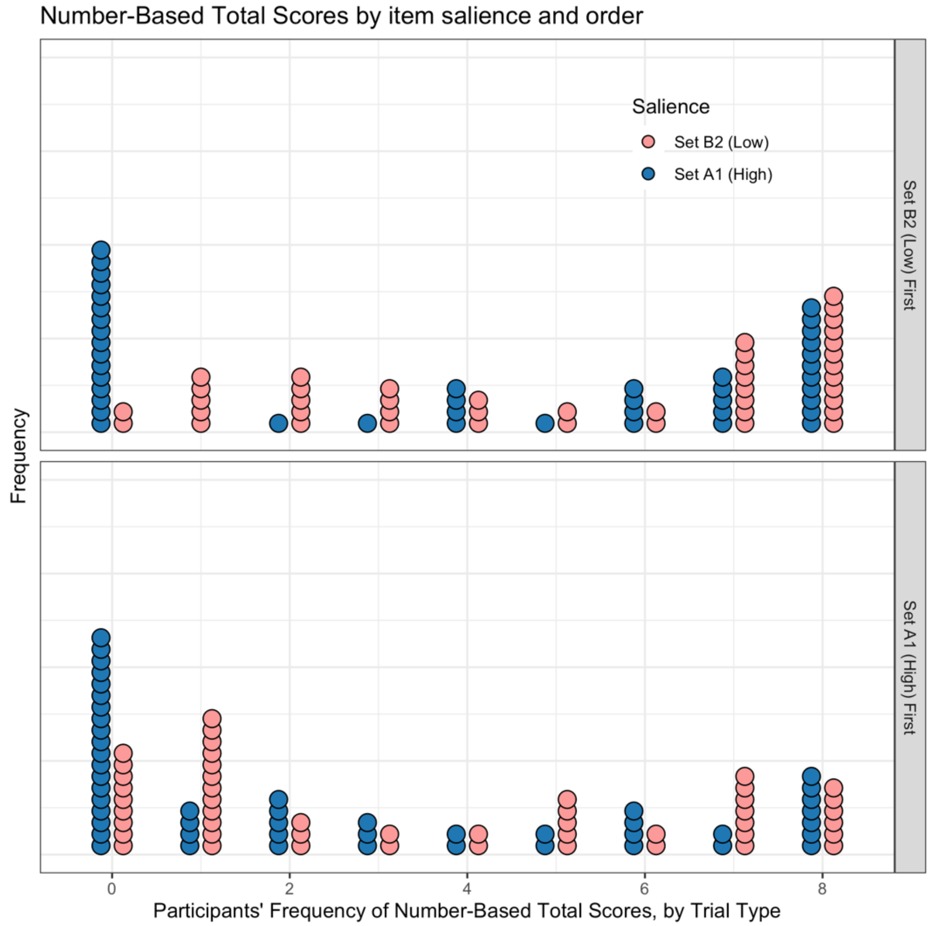

Figure 3. Dot plot histograms of the total frequency of adult participants number-based total score (best match plus revisit trials), broken down by the Order in which the Set B2 trials (low salience condition) appeared (separated between the top vs. bottom panel) and by Item Salience (separated by color). The figure demonstrates the relative differences in frequency of very low number-based total scores (left side of figure) when high vs. low salience trials were viewed first, and the relative differences in frequency of high number-based total scores (right side of figure) when low vs. high salience trials were viewed first.

Figure 4. Crosstabs of frequency of total scores among adult participants, within each bivariate combination of total scores. For example, the ‘5 in the upper left of the left panel represents that among participants who received low salience trials first, 5 selected a number-based match once out of eight low-salience trials and never out of eight high-salience trials. Cells are color-coded by statistical significance-based classifications, with orange (upper left) representing the 38 individuals who rarely chose number (16 received low salience trials first, 22 received high salience trials first); purple (upper right) representing those seven individuals who chose number-based matches more frequently in low- versus high-salience contexts (2 low first, 5 high first); green (lower right) representing the 27 individuals who frequently selected number-based matches regardless of context, i.e., low or high salience conditions (17 low, 10 high); and grey representing observed bivariate combinations that were not categorized (8 low, 11 high). White cells represent bivariate combinations with no observations across either order.