

Citation: Baer, C., Gill, I. K., & Odic, D. (2018). A Domain-General Sense of Confidence in Children. *Open Mind: Discoveries in Cognitive Science*. Advance publication. https://doi.org/10.1162/opmi_a_00020

DOI:
https://doi.org/10.1162/opmi_a_00020

Supplemental Materials:
https://doi.org/10.1162/opmi_a_00020

Received: 11 March 2018
Accepted: 28 September 2018

Competing Interests: The authors declare no conflict of interest in this work.

Corresponding Author:
Carolyn Baer
cebaer@psych.ubc.ca

Copyright: © 2018
Massachusetts Institute of Technology
Published under a Creative Commons
Attribution 4.0 International
(CC BY 4.0) license



A Domain-General Sense of Confidence in Children

Carolyn Baer ¹, Inderpreet K. Gill¹, and Darko Odic¹

¹Department of Psychology, University of British Columbia

Keywords: confidence, certainty, metacognition, approximate number system, development

ABSTRACT

Our minds constantly evaluate the confidence in what we see, think, and remember. Previous work suggests that confidence is a domain-general currency in adulthood, unifying otherwise independent sensory and perceptual representations. Here, we test whether children also possess a domain-general sense of confidence over otherwise independent perceptual dimensions. Six- to 9-year-olds completed either three simple perceptual discrimination tasks—a number task (“Which group has more dots?”), an area task (“Which blob is bigger?”), and an emotions task (“Which face is happier?”)—or three relative confidence tasks, selecting which of two trials they are more confident on. We find that while children’s discrimination performance across the three tasks was independent and constituted three separate factors, children’s confidence in each of three dimensions was strongly correlated and constituted only a single factor. Our results suggest that confidence is a domain-general currency even in childhood, providing a mechanism by which disparate perceptual representations could be integrated.

INTRODUCTION

To learn, represent, and think about the world, our minds constantly deal with uncertainty: Is my friend angry or surprised? Is it safe to cross the street? To make these decisions, we represent and reason about our confidence—the subjective probability of an outcome (Mamassian, 2016)—integrating it across different moments, contexts, and timescales. Confidence representations in turn guide not only our explicit decisions, but even more automatic, perceptual ones: when struggling to follow a conversation at a loud cocktail party, for example, our eyes naturally look at a person’s moving mouth to help decode what our ears can’t hear (McGurk & MacDonald, 1976). What is the origin and nature of confidence representations, and how are they used to compare and integrate information across distinct and independent perceptual domains?

Recent work with adults has suggested that perceptual confidence may operate as a common, domain-general currency, interfacing and integrating across otherwise independent representations to guide optimal decision making. For example, adults show a strong correlation in their ability to judge confidence for line orientations versus spatial frequency, even though these two decisions are behaviorally and neurally independent (De Gardelle & Mamassian, 2014). In childhood, a domain-general sense of confidence would help in part explain how children integrate and compare information across distinct and independent sources. For example, a child trying to decide which of two groups is more socially dominant might compare their confidence in the numerical size of each group against their confidence in the physical size of each group (Pun, Birch, & Baron, 2016).

But, while a domain-general sense of confidence is an appealing mechanism, existing work has not found a correlation in children's confidence across independent perceptual dimensions. Vo, Li, Kornell, Pouget, and Cantlon (2014), for example, measured 5- to 8-year-old children's confidence by having them bet on getting a question right or wrong, and found that children's betting in a simple number perception game (e.g., which box has more dots) and their betting on an emotion recognition game (e.g., which face looks happier) did not correlate. However, tasks requiring children to explicitly rate or gamble on their confidence have long been known to conflate two key components of children's confidence representations: their sensitivity to confidence (the ability to differentiate between states of confidence; e.g., Salles, Ais, Semelman, Sigman, & Calero, 2016), and their response biases, including a general tendency to be overconfident (Butterfield, Nelson, & Peck, 1988; Lipowski, Merriman, & Dunlosky, 2013; Nelson & Narens, 1980). While both components contribute to confidence judgments, response biases across tasks (e.g., overconfidence in one domain and underconfidence in another, as reported by Vo et al., 2014) might influence the detection of underlying similarities in confidence sensitivity. To eliminate these response biases and focus uniquely on sensitivity to confidence, we adapted a method used in the adult visual perception literature that measures participants' *confidence acuity* independent of their response biases by asking them to decide on which of two trials they are relatively more confident, rather than rate how confident they are on any single trial (Baer & Odic, 2018; De Gardelle & Mamassian, 2014).

Here, we use this novel and accessible measure to test children's confidence acuity in three distinct perceptual dimensions: number, area, and emotion perception (Odic, 2018; Vo et al., 2014). We first confirm the domain-specificity of these dimensions in children aged 6 to 9, then investigate whether confidence sensitivity in these dimensions nonetheless correlates, signaling domain-general, or if children's confidence representations are domain-specific before formal schooling and become domain-general as they develop.

METHOD

Participants

Eighty-one 6- to 9-year-old children ($M = 7;11$ [year; months], range = 6;0–10;0, 42 girls) participated in the study, meeting our a priori goal of 40 children per condition and therefore allowing for adequate psychophysical model fits (Halberda & Feigenson, 2008). Three additional children participated but were removed from the sample because they failed to complete at least 90% of the trials. Participants were tested in a quiet room at an on-campus lab or in a quiet area of their schools in Vancouver, British Columbia. All children spoke English and most came from middle-class families.

Materials and Procedure

Children saw custom-made stimuli on an 11.3 in. Apple MacBook Air laptop using Psychtoolbox-3 (Brainard, 1997) scripts, which are available online for free use at <http://odic.psych.ubc.ca/scripts/domaingeneralconfidence.zip>. Children saw three types of stimuli, described in detail below and shown in Figure 1: blue and yellow dots (Number), blue and yellow blobs (Area), and two emotional expressions (Emotion). Children randomly assigned to the Confidence condition ($n = 40$) were asked to reason about their relative confidence in answering two questions, while children in the Discrimination condition ($n = 41$) were simply asked to answer the questions. The Discrimination condition therefore allowed us to confirm the domain-specificity of the perceptual discriminations in these three dimensions, as well as

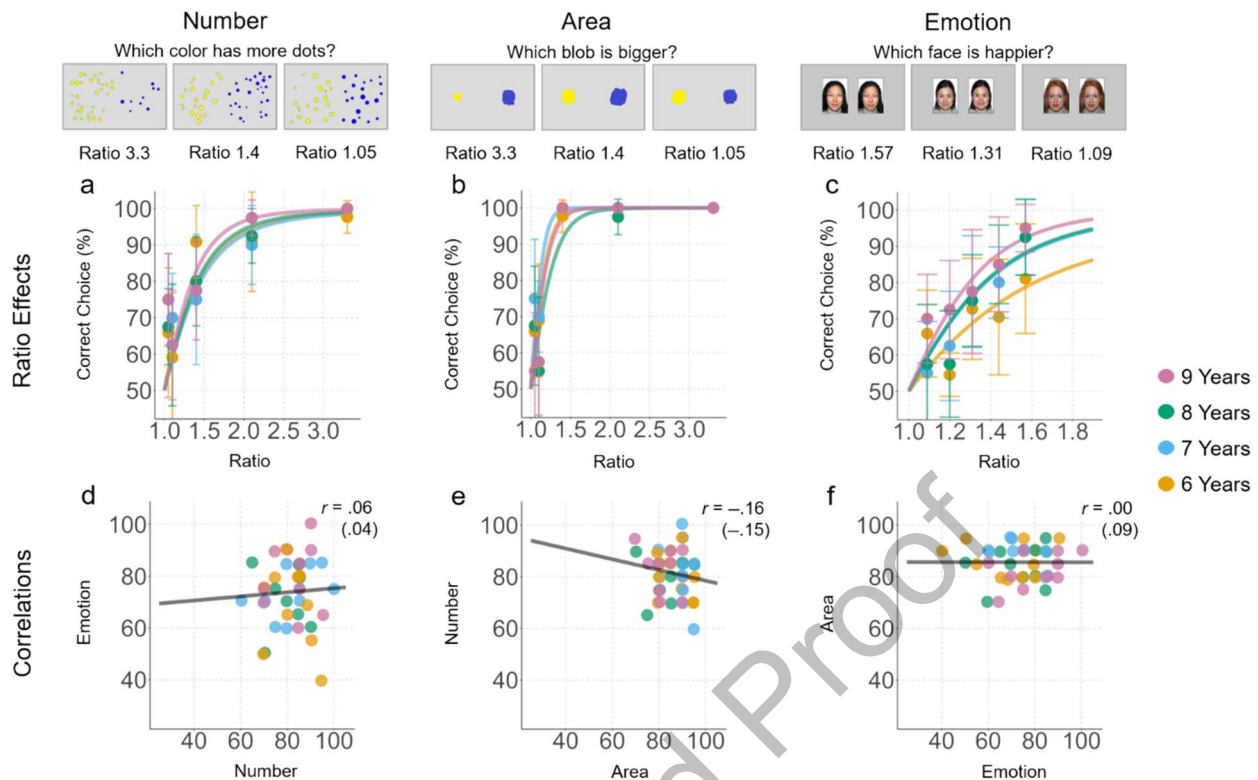


Figure 1. Stimuli and results for the Discrimination condition. Children were shown randomly intermixed trials from the three tasks at one of 5 ratios. As shown in panels a–c, accuracy improved as ratios increased (error bars represent 95% CI). Panels d–f show the correlations between dimensions. Correlations in brackets are controlling for age. No correlations here are significant.

control for the possibility that correlations between dimensions in the Confidence condition are due to other domain-general comparison or task-comprehension abilities.

Discrimination Condition. In the Discrimination condition, children saw a Number, Area, or Emotion trial in a random, intermixed order. This both prevented order effects between the three stimuli types and made the task more interesting for children. To remove the influence of their developing motor skills and inhibitory control, children were asked to either verbalize their answer or point to one side of the screen, and the experimenter pushed a corresponding button. Children received feedback after each trial in the form of a prerecorded female voice that would either give positive feedback (e.g., “That’s right!”) or negative feedback (“Oh no, that’s not right”). Occasionally, the experimenter would give additional feedback to encourage the child to stay engaged in the task (e.g., “That’s okay, let’s do another one”). After completing 12 practice trials that familiarized children with each dimension (4 per dimension), children completed a total of 60 trials (20 per dimension).

Number Discrimination. This task was modeled after dozens of studies exploring children’s approximate number system (ANS), an early sense of number that is broadly shared with other, nonhuman animals (Halberda & Feigenson, 2008). Children saw a set of yellow and blue dots, with the yellow dots on the left and the blue dots on the right (Figure 1) for 1,000 ms—preventing them from counting—and were asked to identify “which side has more dots.” We varied difficulty by manipulating the ratio of blue to yellow dots, showing children one of five ratios on each trial: 3.3 (e.g., 33 yellow dots and 10 blue dots), 2.1, 1.4, 1.1, and 1.05.

Area Discrimination. This task was modeled after studies exploring children's early ability to discriminate area (Odic, 2018). Children were shown a yellow amorphous blob on the left and a blue amorphous blob on the right (Figure 1) for 1,000 ms, and were asked to identify "which blob is bigger." We varied difficulty by manipulating the ratio of pixels in the blue and yellow blobs, showing children one of five ratios on each trial: 3.3 (e.g., 119,130 yellow pixels and 36,100 blue pixels), 2.1, 1.4, 1.1, or 1.05.

Emotion Discrimination. This task was modeled after studies exploring children's emotion discrimination (Vo et al., 2014). Children saw two female faces that differed in emotion on a spectrum from happy to angry for 1,000 ms (Figure 1), and were asked to identify "which face is happier." To generate the stimuli, we took a 100% happy and a 100% angry face from four different female models—two Caucasian and two Asian—and blended the two faces using the FantaMorph software (Abrosoft, 2007). We blended faces in 6.67% intervals, creating eight total blends varying from 100% happy (i.e., 0% angry), through 53.3% happy (i.e., 46.7% angry). We varied difficulty by presenting two faces whose difference was either easy to tell apart (e.g., 93.3% happy vs. 60% happy, a ratio of 1.56) and some that are very difficult (e.g., 73.3% happy vs. 66.7% happy, a ratio of 1.1). This resulted in five different binned ratios: 1.09, 1.2, 1.31, 1.43, and 1.57.

Confidence Condition. This task used identical stimuli to the Discrimination condition, with one simple, but major, change: rather than showing children a single trial and asking them to choose the correct answer, we presented two trials simultaneously and asked children to choose which of two trials they *wanted* to answer (Baer & Odic, 2018; De Gardelle & Mamassian, 2014; Figure 2). Because we rewarded children for their accuracy through positive feedback, we expected that children would maximize their chances of success by choosing the more certain (i.e., easier) question. By varying the difference in difficulty between the two trials, we can identify children who can tell apart only large differences between their confidence (e.g., the difference between "very sure" and "not sure") versus children who can tell apart even small differences in their internal confidence (e.g., between "very sure" and "somewhat sure"), giving us a measure of individual differences to compare across domains. This *relative* confidence task is extensively used in the adult perception literature, as it eliminates the possibility of response biases (e.g., saying "very confident" on every trial).

After 12 practice discrimination-only trials, evenly distributed between each dimension, children completed 45 confidence trials (15 per dimension). On each trial, children were presented with a pair of stimuli made up of either dots (Number), blobs (Area), or faces (Emotion) trials used in the Discrimination condition (note that children only saw pairs of stimuli from a single dimension at a time). As in the Discrimination condition, these stimuli were presented in a random, intermixed order. Children were asked to identify "which of these two questions would you like to do." The trial would stay on the screen until children responded by verbalizing or pointing to their answer, and the experimenter pushed a corresponding button. To keep children motivated to choose the easier question, the selected trial would then expand to fill the screen, and children answered the question as in the Discrimination condition (e.g., judging which side has more dots in the case of Number). After choosing the answer for the selected discrimination trial, children received feedback in the form of a prerecorded female voice. As in the Discrimination condition, the experimenter would occasionally provide additional feedback to encourage the child to stay engaged in the task (e.g., "That's okay, let's do another one!"), but the child *never* received feedback on whether they had successfully selected the easier question (see Smith, Beran, Couchman, & Coutinho, 2008).

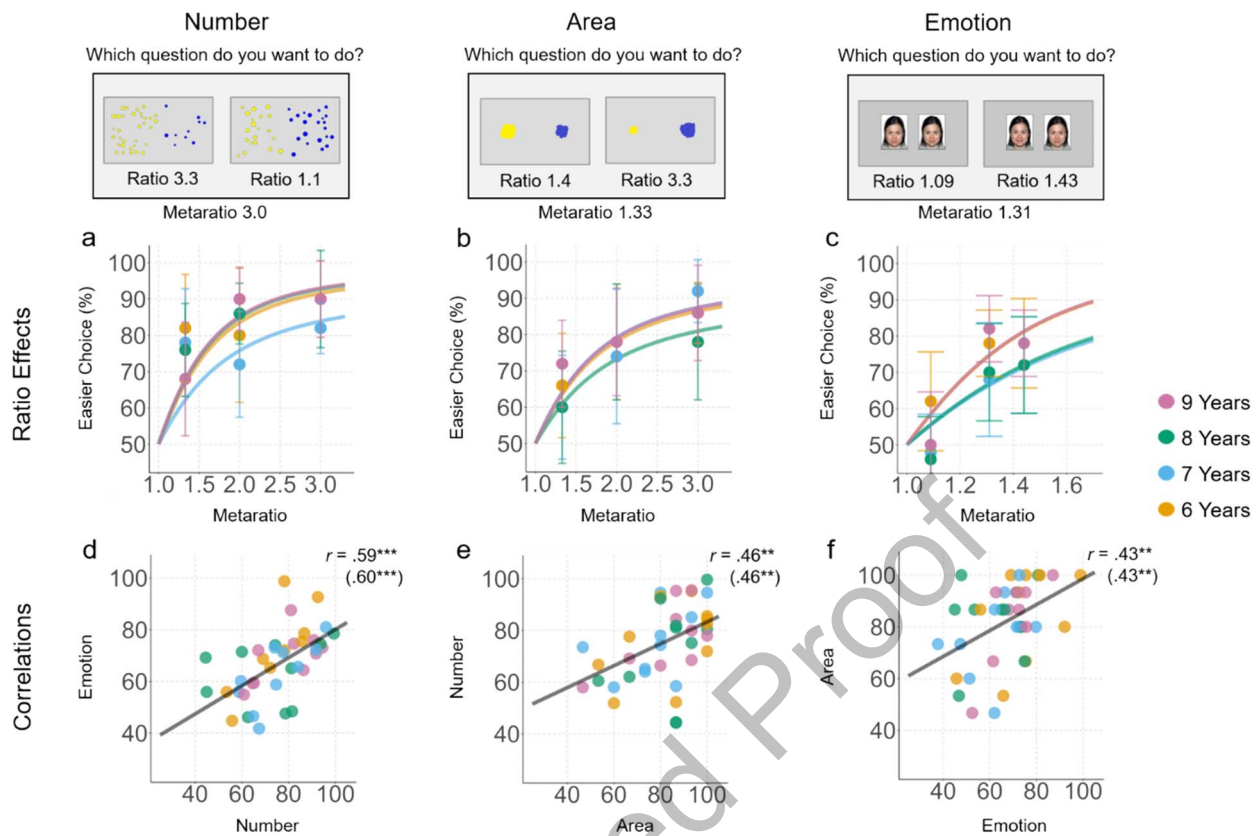


Figure 2. Stimuli and results for the Confidence condition. Children were shown randomly intermixed pairs of trials from one of the three tasks and asked which question they wanted to answer. Then, the screen zoomed in on their selected question and they answered it, just as in the Discrimination trials. As shown in panels a–c, accuracy improved as ratios increased (error bars represent 95% CI). Panels d–f show the correlations between dimensions. Correlations in brackets are controlling for age. ** denotes $p < .01$, *** denotes $p < .001$.

To vary the difficulty and estimate individual differences in the precision with which children could tell apart levels of confidence, we varied the difference in the relative difficulty between the two presented trials (i.e., the “metaratio”—the larger numerical ratio divided by the smaller one). For example, children were shown one Number trial with a ratio of 3.3 on the left (e.g., 33 yellow to 10 blue dots), and a ratio of 1.1 on the right (e.g., 22 yellow to 20 blue dots), yielding a metaratio of 3.0 (3.3 / 1.1). The difference in difficulty between the two trials becomes harder to detect as the metaratio approaches 1.0, much like the difficulty in telling apart two quantities becomes harder to detect as the ratio approaches 1.0. Children were presented with three metaratios per dimension: 3.0, 2.0, and 1.33 (for Number and Area), and 1.44, 1.31, and 1.1 (for Emotion; e.g., an easy 1.57 ratio vs. a hard 1.09 ratio yields a metaratio of 1.44).

In previous research, we have shown that children generally want to maximize their chance of success in relative confidence tasks and therefore choose the trial in which they have more confidence, producing a metaratio effect: the larger the difference between two presented ratios, the more likely children are to indicate the easier trial as the one they are more confident in (Baer & Odic, 2018). However—we also found that, given that we used wording that uses simple vocabulary and avoids advanced mentalistic terms (i.e., “which trial do you *want* to do,” rather than “which trial are you more *confident* on”), some of the children in our sample had significant *below chance* performance in which they consistently selected

the harder of the two trials. Indeed, many of these children would subsequently tell us that they chose the harder trials in order to challenge themselves. While these children's consistent preference for the harder ratios clearly demonstrates the ability to differentiate between their confidence states (which is what we are ultimately interested in), their data also create a bimodal accuracy distribution, leading to violations of several statistical assumptions. For this reason, all children's data were fit both by a psychophysical model assuming that children select the easier of the two trials (see Results), and an inverted version of the same model whereby we assume that children select the harder of the two trials. By using model comparison, we identified 10 children whose data clearly show a preference for harder trials: 3 children showed this behavior across all three dimensions, while 6 children showed this behavior on only one of the three dimensions. For these children, we treated their inverted model as the dependent variable (e.g., a child with accuracy of 10% was modeled to have accuracy of 90%). Although our conclusions remain the same with these 10 children removed or modeled with the noninverted model, we also include a complete report of our analyses with only the noninverted model in the Supplemental Materials (Baer, Gill, & Odic, 2018).

RESULTS

We first report results from the Discrimination condition, which serves as a control condition, allowing us to detect any preexisting correlations between perceptual representations or task understanding in number, area, and emotion perception. Subsequently, we conduct identical analyses on the Confidence condition to see whether confidence acuity correlates across the three dimensions.

Discrimination Condition

Children in the Discrimination condition performed above chance for all three dimensions (see Table 1 for means and tests against chance), and performed significantly better at Area compared to Number (replicating Odic, 2018), and at Number compared to Emotion, $F(2, 80) = 14.67, p < .001, \eta_p^2 = .27$. We also found that children's accuracy on the Emotion trials increased with age, $r(39) = .32, p = .04$, but found no age effects for the Area or Number trials, $r_{Num}(39) = .07, p = .685, r_{Area}(39) = -.26, p = .099$, most likely due to our truncated age range compared to past research in these areas (e.g., Odic, 2018). Additionally, children's accuracy varied as a function of ratio in each of the three dimensions (Figure 1a-c): Number: $F(3.23, 129.01) = 24.66, p < .001, \eta_p^2 = .38$; Area: $F(1.95, 78.16) = 59.57, p < .001, \eta_p^2 = .60$; Emotion: $F(3.67, 146.61) = 16.45, p < .001, \eta_p^2 = .29$.

Table 1. Descriptive statistics, tests against chance, and average estimates of fit to the Weber model.

Dimension	M [95% CI]	t	p	d	# Fit	w [95% CI]	Lapse rate [95% CI]
Discrimination condition							
Number	81.44 [78.57, 84.30]	22.18	<.001	3.46	41	.23 [.17, .30]	.02 [.00, .03]
Area	85.58 [83.42, 87.74]	33.30	<.001	5.20	41	.10 [.10, .10]	.01 [.00, .03]
Emotion	73.86 [69.83, 77.90]	11.95	<.001	1.87	39	.30 [.16, .44]	.08 [.02, .14]
Confidence condition							
Number	75.67 [71.05, 80.29]	11.24	<.001	1.78	37	.53 [.35, .71]	.11 [.08, .17]
Area	82.00 [77.00, 87.00]	12.95	<.001	2.05	37	.32 [.16, .48]	.13 [.07, .19]
Emotion	67.33 [63.27, 71.40]	8.63	<.001	1.36	36	.31 [.19, .43]	.21 [.13, .28]

To simplify data on each child's accuracy as a function of ratio, we fit children's accuracy data to a standard two-parameter psychophysical model widely used in the literature, yielding an estimate of each child's Weber fraction (w): the underlying precision of their number, area, and emotion representations independent of guessing (Halberda & Feigenson, 2008; Odic, 2018; Pica, Lemer, Izard, & Dehaene, 2004). Weber fractions can also be interpreted as the smallest change in a stimulus that can be reliably detected, thus smaller w estimates indicate better acuity. This model assumes that the underlying representations of number, area, and emotion are normally distributed tuning curves with the single parameter w indexing their standard deviation (i.e., precision; for review, see Halberda, Odic, & Koepke, 2014). On top of this standard assumption, the model fits the lapse rate, which accounts for a constant percentage of trials that participants may have been guessing (e.g., a lapse rate of 0.10 indicates that participants were randomly guessing on 10% of trials, independent of ratio). More formally, Weber fractions and lapse rates were estimated using the equation:

$$Accuracy = (1 - lapse) * \Phi\left(\frac{Ratio - 1}{w * \sqrt{1 + Ratio^2}}\right) + \frac{lapse}{2}$$

where Φ is the Gaussian cumulative distribution function. This model was fit to each participant's data for each task using R's *optim* function (R Core Team, 2018) under the assumption of normally distributed errors, converging on the best-fit parameters by minimizing the negative log-likelihood value. We successfully fit the data of all children except for two in the Emotion Discrimination condition (a typical cut-off of $w = 3$ was used to fit data; Halberda & Feigenson, 2008; Odic, 2018; accuracy data for these children were still included in all accuracy analyses). The estimated w values for the three dimensions are shown in Table 1, and replicate previously established values for children of this age (Halberda & Feigenson, 2008; Odic, 2018).

Finally, we found that children's performance on the three dimensions were independent of each other; how well children did on the Number trials did not correlate with how well they did on Area or Emotion, and vice-versa (Figure 1d-f). This result held for both accuracy and w data, and held when we controlled for the effects of age (see Table 2). Thus, we can conclude, consistent with previous work (Odic, 2018; Vo et al., 2014), there is evidence for domain-specificity in children's number, area, and emotion perception.

Table 2. Correlations between tasks with and without controlling for age.

	Accuracy (r)		w (ρ)	
	Area	Emotion	Area	Emotion
Discrimination condition				
Number	-.16 (-.15)	.06 (.04)	.02 (.02)	.00 (.00)
Area		.00 (.09)		.23 (.29)
Confidence condition				
Number	.46** (.46**)	.59*** (.60***)	.34^ (.34^)	.66** (.66**)
Area		.43** (.43**)		.20 (.20)

Note. Accuracy data are correlated using Pearson's r , while model fit estimates are correlated using Spearman's ρ . Correlations controlling for age are shown in brackets. ^ $p < .10$, ** $p < .01$, *** $p < .001$.

Confidence Condition

Children in the Confidence condition also performed above chance for all three dimensions (see Table 1 for means and tests against chance), choosing the easier of the two trials on 75% of trials (95% CI[71.44, 78.78], $t(39) = 13.84$, $p < .001$, $d = 2.19$), suggesting that they reasoned about their relative confidence in the two questions. As in the case of Discrimination, we found that children were best on the Area trials and worst on the Emotion trials, $F(2,78) = 20.50$, $p < .001$, $\eta_p^2 = .35$. And, much as in the Discrimination condition, we found no correlations between accuracy and age, all $r < .09$, potentially due to our restricted range.

Replicating past work using this measure of confidence with Number stimuli, we found that children were more likely to choose the easier question when the metaratio was higher, $F(2,78) = 10.00$, $p < .001$, $\eta_p^2 = .20$ (Baer & Odic, 2018). In other words, children's confidence discrimination was itself ratio-dependent. But, critically, we also found the same metaratio effect for Area, $F(1.53, 59.70) = 5.19$, $p = .014$, $\eta_p^2 = .12$, and Emotion, $F(2,78) = 18.57$, $p < .001$, $\eta_p^2 = .32$ (see Figure 2a–c), suggesting that it is not merely an effect of number perception, and suggesting that this confidence task can be successfully and reasonably used across a variety of stimuli types.

Because performance in the Confidence condition was metaratio-dependent, we fit children's confidence data to the same psychophysical model as the one used in the Discrimination condition, estimating each child's *confidence acuity* (the precision with which they can distinguish their internal confidence states) separately from their guessing behavior. We successfully fit all but eight children on all three tasks and all but one child on at least two tasks using the same criteria of $w < 3$ as in the Discrimination condition, but retain all children's accuracy data for all subsequent analyses. The fit w data are presented in Table 1.

Finally, and most importantly, we found strong correlations between Number, Area, and Emotion confidence discrimination for accuracy, and slightly weaker correlations with w (Figure 2, see Table 2). This result stands in strong contrast to the Discrimination condition and suggests an important degree of domain-generality in confidence perception that is not present when children are merely discriminating each dimension.

Principal Component Analyses. To further confirm the domain-generality of children's confidence perception, we ran two principal component analyses (PCAs), which attempt to simplify a set of variables into factors that explain the maximum possible variance (Hair, Black, Babin, & Anderson, 2009): one for accuracy and Weber estimates for all three dimensions in the Discrimination condition, and one accuracy and Weber estimates for all three dimensions in the Confidence condition.

In the Discrimination condition, there were three components identified in the scree plot and associated eigenvalues, clustered by dimension. To improve interpretability, the factor loadings (i.e., the correlations between variables and the extracted components) were varimax-rotated. Number, Area, and Emotion each uniquely mapped onto separate components, consistent with the interpretation that each dimension is independent (see Table 3 for factor loadings).

In contrast, only one component was identified in the Confidence condition, consistent with a domain-general system. Factor loadings are shown in Table 3 (because only one component was extracted, these could not be varimax-rotated). An additional analysis in the Supplemental Materials (Baer et al., 2018) also shows that the two PCAs extracted significantly different proportions of variance. In sum, despite strong evidence that the underlying

Table 3. Factor loadings of measures by dimension.

Measure	Discrimination condition			Confidence condition
	Component 1	Component 2	Component 3	Component 1
Number				
Accuracy	.126	.953	.009	-.804
w	.046	-.961	.037	.958
Area				
Accuracy	-.966	-.048	-.047	-.820
w	.971	.028	.073	.865
Emotion				
Accuracy	-.146	.186	-.871	-.841
w	-.023	.146	.900	.908
Eigenvalue	2.07	1.87	1.44	4.52
Variance Explained	35%	31%	24%	75%

Note. Lower w values indicate better precision.

perceptual discriminations are domain-specific, confidence discriminations are domain-general from at least the age of 6.

GENERAL DISCUSSION

Our data are the first to show evidence of domain-generality in 6- to 9-year-old children's sense of confidence: while children's perceptual discrimination of number, area, and emotion were dissociated, their *confidence* judgements over these same dimensions are strongly correlated and constitute a single factor, extending previous work in adults (De Gardelle, Le Corre, & Mamassian, 2016; De Gardelle & Mamassian, 2014). We find, therefore, that children as young as age 6 share a domain-general sense of confidence with adults, suggesting that confidence is either domain-general throughout development, or else is combined before children begin formal schooling.

Our results hold several implications about the nature and origin of confidence representations.

First, a domain-general sense of confidence should allow children and adults to compare information *across* perceptual boundaries. Under many models, and consistent with our discrimination data, perceptual magnitudes are represented on distinct scales (e.g., Odic, 2018), making cross-magnitude comparison difficult. A domain-general sense of confidence could, therefore, act as a universal translator between magnitudes: given that confidence in each dimension could be represented on a scale that is shared broadly across all magnitudes, observers should be able to easily compare and decide which information is most reliable in a given context. For example, if our friend says a word that sounds like "noodle" while talking to a friend about dogs, we can use a domain-general sense of confidence to compare the auditory cues to the social cues and determine that they must have said "poodle." Similarly, an observer faced with a spontaneous discrimination task in which number is easier to discriminate than area should prioritize numerical information over other magnitudes (e.g., Cantlon, Safford, & Brannon, 2010).

Second, our results suggest the domain-general confidence scale is itself subject to individual differences and that any intervention that helps teach an individual how to make more

precise confidence decisions should impact confidence across the board, helping or hindering children's confidence across perceptual magnitudes. Furthermore, because we find that confidence precision is correlated even in children, these findings support the view that confidence itself is *represented* in a domain-general format (De Gardelle et al., 2016; De Gardelle & Mamassian, 2014), though future work will need to uncover whether this pattern holds at even younger ages.

In conclusion, we find that children as young as 6—much like adults—have a domain-general sense of confidence that crosses otherwise independent perceptual representations. Our work places confidence in a broader developmental context and shows continuity between adults and developing children's minds.

ACKNOWLEDGMENTS

We would like to acknowledge the support of the families and schools that participated in this project.

FUNDING INFORMATION

Carolyn Baer, Social Sciences and Humanities Research Council of Canada (<http://dx.doi.org/10.13039/501100000155>), Canada Graduate Scholarship—Doctoral. Darko Odic, Social Sciences and Humanities Research Council of Canada (<http://dx.doi.org/10.13039/501100000155>), Insight Development Grant.

AUTHOR CONTRIBUTIONS

Carolyn Baer: Conceptualization: Lead; Data curation: Equal; Formal analysis: Lead; Funding acquisition: Supporting; Investigation: Supporting; Methodology: Equal; Project administration: Lead; Resources: Equal; Software: Supporting; Visualization: Lead; Writing – original draft: Lead; Writing – review & editing: Equal. Inderpreet K. Gill: Investigation: Lead; Methodology: Equal; Resources: Equal; Writing – review & editing: Equal. Darko Odic: Conceptualization: Supporting; Data curation: Equal; Formal analysis: Supporting; Funding acquisition: Lead; Methodology: Equal; Resources: Lead; Software: Lead; Supervision: Lead; Visualization: Supporting; Writing – original draft: Supporting; Writing – review & editing: Equal.

REFERENCES

- Abrosoft. (2007). FantaMorph, version 4 [Computer software]. Beijing, China: Author. Retrieved from <http://www.fantamorph.com>
- Baer, C., Gill, I. K., & Odic, D. (2018). Supplemental material for “A domain-general sense of confidence in children.” *Open Mind: Discoveries in Cognitive Science*, 1(4). doi:10.1162/opmi_a_00020
- Baer, C., & Odic, D. (2018). *Sensitivity to certainty undergoes unique development in childhood*. Manuscript submitted for publication.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436. doi:10.1163/156856897X00357
- Butterfield, E. C., Nelson, T. O., & Peck, V. (1988). Developmental aspects of the feeling of knowing. *Developmental Psychology*, 24, 654–663. doi:10.1037/0012-1649.24.5.654
- Cantlon, J. F., Safford, K. E., & Brannon, E. M. (2010). Spontaneous analog number representations in 3-year-old children. *Developmental Science*, 13, 289–297. doi:10.1111/j.1467-7687.2009.00887.x
- De Gardelle, V., Le Corre, F., & Mamassian, P. (2016). Confidence as a common currency between vision and audition. *PLOS ONE*, 11, e0147901. doi:10.1371/journal.pone.0147901
- De Gardelle, V., & Mamassian, P. (2014). Does confidence use a common currency across two visual tasks? *Psychological Science*, 25, 1286–1288. doi:10.1177/0956797614528956
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Pearson.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “number sense”: The approximate number system

- in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457–1465. doi:10.1037/a0012682
- Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. In D. C. Geary, D. Berch, & K. Koepke (Eds.), *Evolutionary origins and early development of number processing* (pp. 305–333). San Diego, CA: Academic Press.
- Lipowski, S. L., Merriman, W. E., & Dunlosky, J. (2013). Preschoolers can make highly accurate judgments of learning. *Developmental Psychology*, 49, 1505–1516. doi:10.1037/a0030614
- Mamassian, P. (2016). Visual confidence. *Annual Review of Vision Science*, 2, 459–481. doi:10.1146/annurev-vision-111815-114630
- McCurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. doi:10.1038/264746a0
- Nelson, T. O., & Narens, L. (1980). A new technique for investigating the feeling of knowing. *Acta Psychologica*, 46, 69–80. doi:10.1016/0001-6918(80)90060-8
- Odic, D. (2018). Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental Science*, 21, e12533. doi:10.1111/desc.12533
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306, 499–503. doi:10.1126/science.1102085
- Pun, A., Birch, S. A. J., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences*, 113, 2376–2381. doi:10.1073/pnas.1514879113
- R Core Team. (2018). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Salles, A., Ais, J., Semelman, M., Sigman, M., & Calero, C. I. (2016). The metacognitive abilities of children and adults. *Cognitive Development*, 40, 101–110. doi:10.1016/j.cogdev.2016.08.009
- Smith, J. D., Beran, M. J., Couchman, J. J., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin & Review*, 15, 679–691. doi:10.3758/PBR.15.4.679
- Vo, V. A., Li, R., Kornell, N., Pouget, A., & Cantlon, J. F. (2014). Young children bet on their numerical skills metacognition in the numerical domain. *Psychological Science*, 25, 1712–1721. doi:10.1177/0956797614538458

Corrected Proof