Précis of *Developing a Sense of Certainty*

Carolyn Baer
Department of Psychology, University of British Columbia
December, 2020

We get information about the world from many imperfect sources: hearing muffled voices, dream-like memories, and misinformed testimony from peers. How can we make sound decisions amid all these flawed sources? Adults have a sophisticated cognitive toolbox of experience to navigate the uncertain landscape we face. Children, however, don't have years of experience to form accurate expectations (after all, they don't even have fully developed perceptual and cognitive systems). And yet, children grow up to be adults who possess a sophisticated cognitive toolbox that can generally tell truth from falsehoods.

This dissertation proposes that children develop this toolbox, in part, by fine-tuning a sense of *confidence*: subjective assessments of how likely we are to be correct about our thoughts, decisions, and knowledge. Confidence prompts us to either trust or doubt our own judgement, signalling whether we know enough to act or whether we have missed something and should be skeptical.

In seven studies, I test longstanding questions from psychology, education, neuroscience, and epistemology about what confidence is and where it comes from. I develop a new method for assessing confidence in children that eliminates previous barriers to studying these questions, yielding new discoveries about what cognitive processes are involved in generating a feeling of confidence. I find that children fine-tune their sense of confidence independently from other skills, and that confidence is a domain-neutral unit allowing for integration across distinct cognitive systems. This work provides critical evidence about which properties of confidence are malleable and which properties are constant throughout development, providing broad implications for educational interventions and theoretical accounts alike.

The Challenges of Studying Confidence

The study of how the mind represents confidence is challenging. Confidence is subjective. It is not a property of events in the world, but a property of the mind – and of that mind's representation *of itself* at that. Confidence also does not occur in isolation; we are always certain *about* a decision or a piece of knowledge. It is therefore challenging to precisely define and measure confidence, let alone understand the computations that generate a feeling of confidence.

Studying confidence in children adds further challenges. Young children famously make unfounded claims like 'I *know* I can fly!', suggesting that their certainty is not well-calibrated to reality. However, this does not mean that the *signal* of confidence they generate is entirely uninformative. Drawing from signal detection theory, there is an important distinction between our ability to report on an internal signal (known as 'bias') and the quality of that signal (known as 'sensitivity'; Green & Swets, 1966). A child might experience a sensible confidence signal but not know how to communicate it to others. One solution to separate these two components is to apply formal decision-making models (e.g., d' and meta-d', AUROC, phi or gamma correlations; Maniscalco & Lau, 2012, 2014; Nelson, 1984; Vo et al., 2014). These models typically require many trials (>100; Fleming, 2017), which is incredibly challenging to achieve with younger populations. Further, many measures of confidence depend on scale use (e.g., "really sure", "kind of sure", "not so sure"; Hembacher & Ghetti, 2014) or mental state terms (e.g., "do you *know* it?"; Rohwer et al., 2012). These entirely exclude very young children and non-verbal populations from the study of confidence simply because of the format of the confidence report.

To study how the mind generates a feeling of confidence, we need a child-friendly measure of the confidence signal separately from the confidence report.

A Forced-Choice Measure of Children's Confidence

This dissertation presents a new measure of children's confidence reasoning. In contrast to existing measures requiring an explicit report (like asking 'are you sure?'), this method asks children to compare their confidence between two options: one designed to elicit high confidence and one low (see Figure 1; Mamassian, 2020). If children are sensitive to the difference in confidence, they should strategically choose the question eliciting higher confidence.

This method bypasses challenges in interpreting or communicating confidence and allows us to directly examine children's sensitivity to confidence states. Children do not need to calibrate their confidence against an arbitrary scale in this method (i.e., deciding which state of confidence is 'enough' to be considered 'high'). Instead, they must compare the two states of confidence *to one another*, without ever needing to label those confidence states. Children who are highly sensitive to confidence can tell apart similar states (e.g., 'very sure' from 'extremely sure') while other children might only tell apart highly disparate states (e.g., 'very sure' from 'not at all sure').

To detect these differences in confidence sensitivity, we need questions designed to elicit a wide spectrum of confidence signals. I turned here to the vast literature on children's developing numerical abilities (Dehaene, 2011; Halberda et al., 2012; Odic & Starr, 2018). Even human infants possess intuitive approximate number sense (Halberda et al., 2012; Halberda & Feigenson, 2008; Izard et al., 2009). This can be easily assessed by asking children which of two groups of dots has 'more dots' (see Figure 1; Halberda & Feigenson, 2008). Importantly, the number sense is ratio-dependent: it is easier to discriminate large ratios (like 10 vs 20 or 100 vs 200) than small ones (like 10 vs 11 or 100 vs 110). Because accuracy varies with ratio, so does confidence (Halberda & Odic, 2014; Vo et al., 2014). We can therefore manipulate the experienced confidence by varying the ratio of two quantities. This gives us a child-friendly task that can be modified to create theoretically infinite degrees of confidence.

This new paradigm for studying children's confidence opens new avenues for future work on the origins and implications of confidence. This paradigm can be made entirely non-verbal for non-verbal human and non-human populations. It also permits precise measurement of exactly *how* confidence impacts learning. For instance, we can test whether it is more important that those response biases drive learners to seek out opportunities, or that the fine-tuning of their sensitivity to confidence permits it to be used more optimally (see Baer & Odic, 2020).

Here, I use this new forced-choice measure of confidence in children to test broad theories about how the mind generates confidence signals. I specifically focus on children between 3-9 years, when reasoning about confidence is thought to change dramatically (e.g., Lockl & Schneider, 2004; Lyons & Ghetti, 2011; O'Leary & Sloutsky, 2017). My approach leverages the change inherent in children's development to uncover which properties of confidence are stable and part of the core structure of confidence. I examine three such properties in this dissertation. First is the input used to make the confidence signal. Second is the generality of the cognitive process computing confidence. Third is the link between generating confidence judgments for the self and for others.

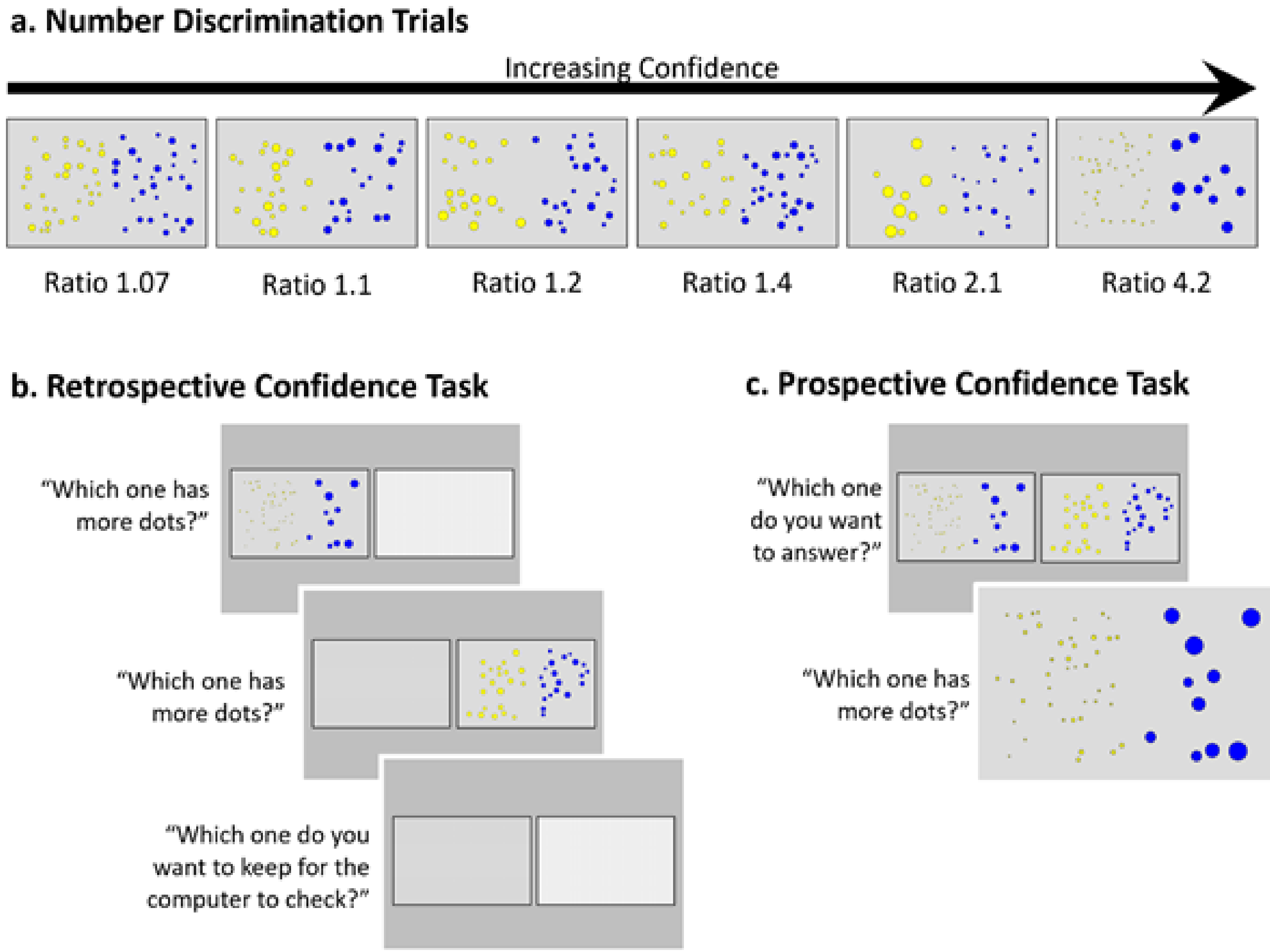

Figure 1: Sample Stimuli Used in the Forced-Choice Confidence Task.

Section (a): Sample number discrimination trials. Children indicate which colour has more dots. Section (b): A retrospective version of the forced-choice confidence task. Children first answer the left question, then right, then select the answer they were most confident in. Section (c): A prospective version. Children first answer the confidence question by selecting the trial they most expect to get correct, then answer only that question.

Confidence in Number Judgments Develops Independently of Numerical Ability

Baer & Odic (2019). Certainty in numerical judgments develops independently of the Approximate Number System. Cognitive Development.

Children's metacognitive abilities improve on most measurements as they age (Ghetti et al., 2013). For instance, 5-year-olds show less consistency in using high and low bets in response to uncertainty than 8-year-olds (Vo et al., 2014). Understanding why this change occurs will point us towards the components of confidence that are stable and part of the core structure of confidence.

Perhaps this development is driven not by improving sensitivity, but by learning how to interpret the confidence signal to report to others. In this case, the input and computation of the confidence signal are stable over time, but existing measures cannot detect this because they are affected by children's improving labelling behaviours. The forced-choice method directly tests this, asking children to compare confidence states to one another rather than explicitly label them. If development is driven entirely by change in interpreting and reporting confidence, there should be no development on the forced-choice measure.

Instead, the new measure also finds development. Across two studies (N = 161 children aged 3-7 years), 3 and 4-year-olds chose the option they were most likely to answer correctly only when the difference was large (akin to 'unsure' versus 'very sure'), while 7-year-olds could tell apart closer differences ('very sure' versus 'extremely sure,' see Figure 2). Children's ability to report their confidence is therefore not the sole mechanism of developmental change in confidence reasoning.

An alternative mechanism for development is that the *input* used to generate a confidence signal is changing. Here, there are two main arguments. First, confidence is always about a decision (e.g., 'I am confident *that there are more blue dots than yellow dots*'). Some theories argue that confidence directly reflects the quality of input used to make those decisions (Galvin et al., 2003; Maniscalco & Lau, 2012; Pouget et al., 2016). If that decision-making capacity fine-tunes over development, as the number sense does (Halberda & Feigenson, 2008; Odic, 2018), then confidence reasoning would *appear* to develop alongside it. I refer to this as the *Direct* account, as it proposes that confidence is computed directly from the same input used to make the decision.

The second argument is that confidence has little to no access to the quality of the decision input (Carruthers, 2009; Koriat, 1993). Instead, confidence is inferred from the outputs of other processes. We use 'cues' like reaction time or anxiety states, noting that we tend to be more accurate (and therefore should be more confident) when we make fast, calm decisions (Koriat & Ackerman, 2010). Confidence could develop under this account because new cues are learned or because children assign new weights to the cues they already attend to. I refer to this as the *Inferential* account given its proposal that confidence is an inference.

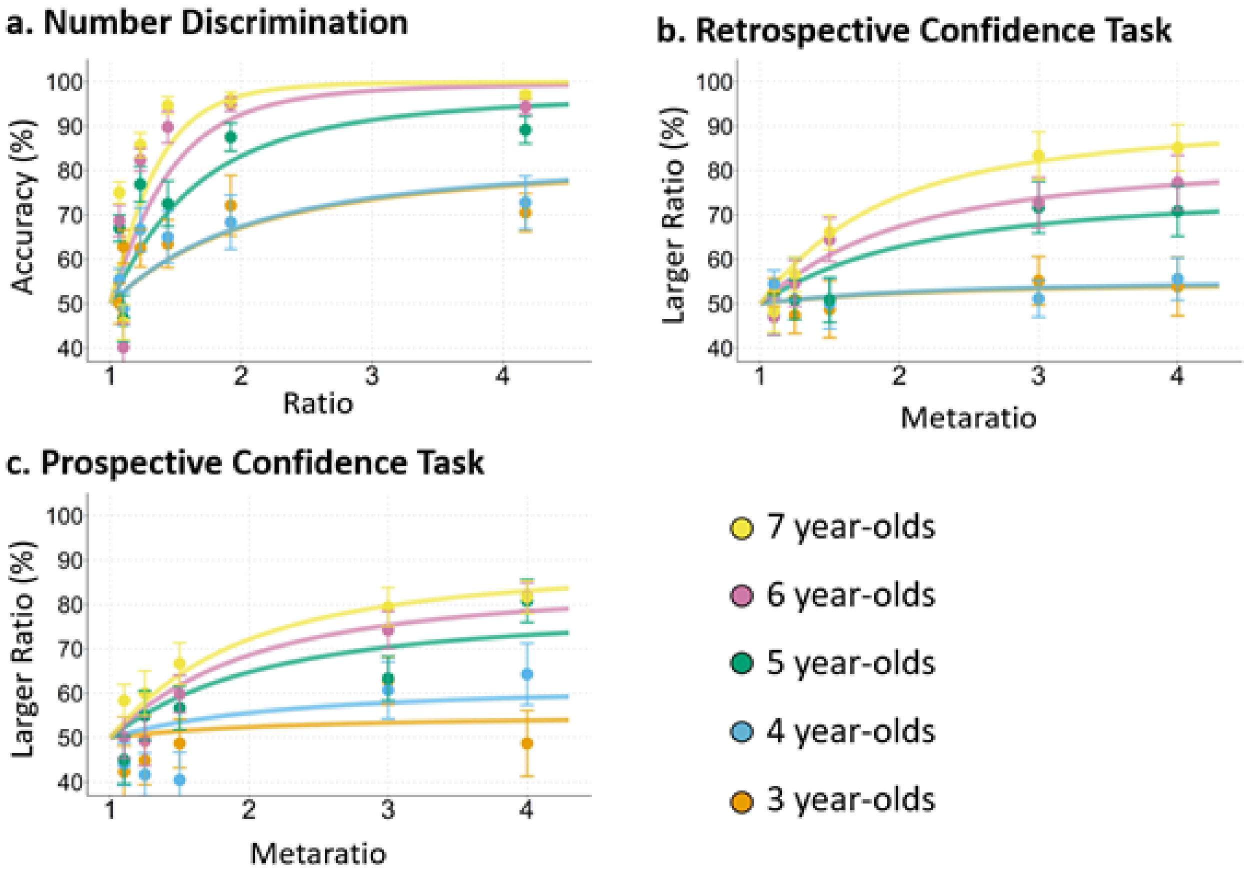

Figure 2: Accuracy by Ratio on the Number Discrimination Trials, and by Metaratio on Confidence Trials.

Section (a): Children chose the more numerous set above chance of 50%. Accuracy was higher both for larger ratios and for older children. Sections (b) and (c): The same patterns, but for performance on the two versions of the confidence task. Children chose the number question with the larger and therefore easier ratio above chance. Choice of the larger ratio was highest for older children and for larger 'metaratios' (the ratio between ratios on the number discrimination questions. Larger metaratios mean the two questions they chose between were more disparate in difficulty). Error bars represent 1 SE, and curves are estimated using a standard psychophysical model (see Odic, 2018).

I test between these two accounts by measuring children's decision-making accuracy and confidence judgments in a number discrimination task. The Direct account predicts that confidence should not develop if decision-making accuracy is controlled for, as both accuracy and confidence are based on the same input: the inherent imprecision of the intuitive number representations. The Inferential account predicts remaining developmental change in confidence sensitivity over and above decision accuracy.

Across both studies, there was remaining developmental change in children's confidence choices when controlling for numerical discrimination accuracy. Confidence reasoning therefore develops in childhood over and above the developing perceptual and cognitive systems themselves. This finding signals additional malleability in confidence than what is accounted for under some theories. One possibility consistent with this data is that confidence is an inferential process and that cues to confidence must be learned or rebalanced over time. Another could be that the computational process of confidence is itself changing, and that children must learn how best to translate their inputs into a sensible confidence signal. Any future theories about confidence should account for this developmental change.

Children's Perceptual Confidence is Domain-General

Baer, C., Gill, I.K., & Odic, D. (2018). A domain-general sense of confidence in children. Open Mind.
Baer, C. & Odic, D. (2020). Children flexibly compare their perceptual certainty within and across perceptual domains. Developmental Psychology.

We experience confidence in a wide range of circumstances, whether that be estimating the items in our grocery basket ("am I *sure* it's less than 8?"), remembering a new colleague's name ("she said it was Sandy, *right*?"), or even evaluating an item's value ("will this cookie *really* make me happy?"). These decisions are made by diverse cognitive processes and exist in independent units (Odic, 2018; Vo et al., 2014). For instance, it seems nonsensical to say that something is 'more 8 than happy' because numbers and emotions are not interchangeable units. An important question for understanding the structure of the mind is whether confidence is similarly computed by diverse cognitive processes or is instead a centralized process using a common currency.

There are mixed findings in the literature about the domain-generality of confidence reasoning in both adults and children. Adults' confidence reasoning is correlated across perceptual tasks, but generally not correlated between perception and memory (Rouault et al., 2018; Vaccaro & Fleming, 2018). This broadly suggests that the format and content of confidence representations is likely distinct for these domains. There is currently no evidence of correlated individual differences across domains in childhood (Bellon et al., 2020; Geurten et al., 2018; Vo et al., 2014). This is true even for the distinct perceptual dimensions like those that correlated in adults (Vo et al., 2014).

Children in these past studies always explicitly rated their confidence. This conflates children's sensitivity in their confidence representations for their response biases, such as their general overconfidence (Fleming & Lau, 2014; Maniscalco & Lau, 2012; Nelson, 1984). While both components contribute to confidence judgments, different 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 reasoning. The forced-choice paradigm is therefore useful here as well to remove the influence of response biases.

In one study (N = 81 6-9-year-olds), we found very high correlations between confidence choices across three distinct perceptual dimensions (Number, Area, and Emotion; see panel (a) of Figure 3). These three dimensions are ideal candidates for testing the domain-generality of confidence, as they are all independent of each other at the perceptual level (e.g., individual differences in area perception do not correlate with number perception; Odic, 2018). Replicating this, we found little-to-no correlation between the three tasks when children were simply asked to answer the perceptual questions (largest r = -.16, *n.s.*). This reiterates that three separate perceptual systems are encoding the low-level information in the stimuli. Hence, any correlations in children's sense of certainty across these three domains could not stem from shared representations or similar developmental paths. In strong contrast, when children evaluated their confidence between two number questions (or area or emotion), we found very high correlations between confidence choices for all three dimensions (smallest $r = .42$, $p < .05$). Participants who could detect fine differences in certainty in the Number task also could detect fine differences in the Area and Emotion tasks, and vice-versa.

Additionally, a domain-general account makes the strong prediction that confidence is a common currency, such as "likelihood of being accurate", between independent tasks (De Gardelle et al., 2016; De Gardelle & Mamassian, 2014). If true, children should be able to compare their certainty states *across* perceptually distinct dimensions just as effectively as they can compare it *within* those dimensions. If not, children should find it difficult if not impossible to compare between the distinct perceptual units (like comparing '8' to 'happy'). Forty-eight 6 and 7-year-olds were presented with pairs of Number, Area, and Emotion perceptual discriminations, but this time the pairs were either Within-Dimension (e.g., number/number) or Across-Dimensions (e.g., area/emotion). Consistent with the common currency account, children could compare their certainty between different perceptual task types, $t(47) = 3.73$, $p < .001$, $d = .54$, which was not different from their performance on comparisons within the same perceptual task type, $F(1, 47) = .075$, $p = .390$, $BF_{10} = .073$, see Figure 3.

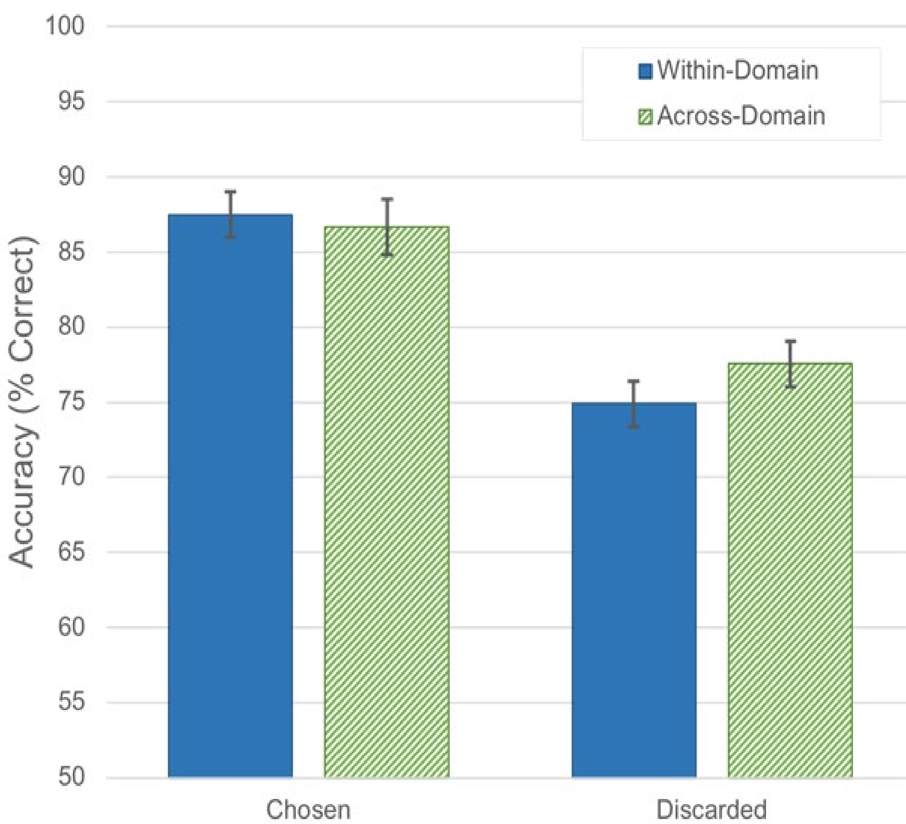

Figure 3: Children's Accuracy on Perceptual Decisions in the Confidence Task.

Results are grouped based on whether the child subsequently chose that question to keep (an indication of high confidence), or to discard (an indication of low confidence). Children were more accurate on items they chose to keep than those they chose to discard, regardless of whether the comparison was within a domain or across perceptual domains. Error bars represent 1 standard error.

Overall, these findings suggest that the ability to evaluate certainty is a domain-neutral currency that bridges across disparate and independent perceptual systems, even in childhood. Building on my earlier work, this is inconsistent with a view of confidence as a direct computation of the decision's quality (which exists in the same units as the decision, like 8 items *plus or minus 2 items*). This work also signals the potential for domain-general transfer of confidence reasoning across perceptual tasks for children. For instance, a recent study found that training young adults' confidence sensitivity using periodic feedback about the accuracy of their confidence judgments led to improved confidence sensitivity on an unrelated task (Carpenter et al., 2019). These results hold potentially powerful implications for educational practices, as metacognitive skills are considered important for effective learning (e.g., Lockl & Schneider, 2004). Our results suggest that similar training effects could occur early as primary school.

These findings also contribute a critical piece to the longstanding puzzle of human consciousness. Some theorists argue that metacognitive reasoning is the hallmark of conscious experience – the very ability that allows us to realize that we are thinking (and therefore 'are'; Dehaene et al., 2017). Evidence of a common currency greatly aids this account, as one prominent view of consciousness argues that a global workspace facilitates our conscious thought (Baars, 1993), but this view lacks a concrete proposal for how domain-specific units could be effectively compared and integrated. If confidence has such a common unit, as the current study suggests may be true even in childhood, then this would be a likely candidate for how conscious thought is achieved (Shea & Frith, 2019).

Confidence Sensitivity Does Not Encompass Social Reasoning

***Baer, C.,** Malik, P., & Odic, D. (2021). Are children's judgments of another's accuracy linked to their metacognitive confidence judgments? Metacognition and Learning*

The findings above support thinking of confidence as a domain-general process with a domain-neutral unit. Some theories of confidence make specific predictions about *how* broadly confidence is computed. The Mindreading account proposes a single process for calculating confidence in one's own knowledge and in the knowledge of others (Carruthers, 2009; Goldman, 2006; Gopnik, 1993). Effectively, children are thought to track the relationship between accurate answers and predictive behavioural cues like response time (which tends to be highly correlated with both accuracy and confidence, Rahnev et al., 2020). Confidence could then be a single process taking these cues and flexibly using them to evaluate one's own or another's likely success (e.g., mindreading).

If confidence is calculated broadly enough to facilitate reasoning about the self and others, then reasoning about one's own confidence and mindreading should correlate. However, the current literature provides conflicting evidence on this relationship. Several studies document correlations and similarities between this kind of self and other reasoning (e.g., Gopnik & Astington, 1988; Kuzyk et al., 2020; Lecce et al., 2015; Lockl & Schneider, 2007; Paulus et al., 2014), while others do not (Bernard et al., 2015; Kim et al., 2020; van Loon & van de Pol, 2019). It thus remains unclear when, if ever, confidence and mindreading should correlate in childhood and what that means for the operation of these two abilities in the mind.

As outlined in both sets of studies above, a major challenge for past confidence studies is the conflation of the underlying sensitivity to confidence and response biases (including understanding mental state language like 'know'). This may be particularly problematic when looking for a correlation with mindreading, which is also tested using mental state verbs (e.g., "know" or "think;" e.g., Wellman & Liu, 2004; Wimmer & Perner, 1983). While the shared language signals that judgments of knowledge in both the self and others contribute to a shared concept of 'knowledge', relying on this language may artificially link the processes leading to such judgments, which may themselves be distinct.

The forced-choice confidence task can offer new insights by removing this additional source of variability from the confidence task. The same forced-choice logic can also be applied to remove the need for mental state language when reasoning about others. A typical mindreading task might involve asking if one agent knows an object's location (Wellman & Liu, 2004). Instead, we can modify this to be a forced choice between *two* individuals who vary in their knowledge, and ask which individual is *more* knowledgeable (known as 'selective social learning;' e.g., Birch et al., 2008; Koenig & Harris, 2005; for review, see Mills, 2013).

In three studies with 4-7-year-olds (N = 304), performance on the forced-choice confidence and selective social learning tasks did not reliably correlate with one another, suggesting that confidence reasoning about the self may not extend to reasoning about others. We used the forced-choice confidence task with area discrimination questions (e.g., "which shape is bigger?") and created a complementary selective social learning task also relying on area discriminations (e.g., "which artist drew the shape closest to the target size?"). Both the confidence and selective social learning tasks replicated key findings of their respective literatures (e.g., a strong preference for the more accurate informant; Einav & Robinson, 2010). Nonetheless, we found no evidence of common processing, as there were no reliable correlations (even through a mega-analysis of the three studies to increase power, $r(303) = .09$, *n.s.*).

These findings are relevant to several broad theories across disciplines. They demonstrate that confidence is not a simple transformation of mindreading abilities, a popular account within epistemology and developmental psychology (Carruthers, 2009; Goldman, 2006; Gopnik, 1993). Instead, these results point to key differences, at the very least in terms of what cues are used for judgments about the self versus others, but possibly also in the cognitive processes that generate those judgments. These results also bear evidence against the recent Cultural Origins hypothesis of metacognition (Heyes et al., 2020). This hypothesis argues that children learn to reason about confidence by watching others model good metacognition or by having a teacher guide them, rather than metacognition emerging either innately through genetic programming or through non-social experience. A critical prediction of this account is that children with good social learning skills should show the best metacognitive skills (see Heyes et al., 2020, p. 358), which we did not find.

The Nature of Confidence Representations

We are constantly bombarded with information and must rely on cognitive tools to sort fact from fiction. This dissertation takes a programmatic and theory-driven approach to explore one such tool that represents the strength of subjective evidence: our sense of confidence. Building on work demonstrating that even children can reason metacognitively (Goupil et al., 2016; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2011), the current work demonstrates an important place for developmental work within the broader research on confidence, teasing apart accounts of how confidence is computed. In accordance with recent theories that confidence is a probabilistic representation of the accuracy of a decision (Meyniel et al., 2015; Pouget et al., 2016), the current evidence of a domain-general unit also gives additional weight to arguments of rationality in childhood (e.g., Gopnik & Bonawitz, 2015; Sobel & Kushnir, 2013). This provides a potential means by which subjective and objective information could be integrated. My ongoing research program further investigates this possibility, (see Baer & Odic, 2021a, 2021b, Baer, Ghetti & Odic, 2021). Thus, far from being subject to the whims of others, children possess a sense of confidence that combines multiple sources in information to create broadly-usable assessments of truth in the world.

References

Baars, B. J. (1993). *A cognitive theory of consciousness*. Cambridge University Press.

Baer, C., & Odic, D. (2020). The relationship between children's approximate number certainty and symbolic mathematics. *Journal of Numerical Cognition*, *6*, 50–65. https://doi.org/10.5964/jnc.v6i1.220

Baer, C., & Odic, D. (2021)a. Mini Managers: Children strategically divide cognitive labor among collaborators, but have a self-serving bias. *Manuscript in preparation*.

Baer, C., & Odic, D. (2021)b. Local, but not Global, Confidence Impacts Children's Strategic Choice. *Manuscript under review for CogSci 2021*.

Baer, C., Ghetti, S., & Odic, D. (2020). Perceptual and Memory Metacognition in Children. *Manuscript under review for CogSci 2021*.

Bellon, E., Fias, W., & Smedt, B. D. (2020). Metacognition across domains: Is the association between arithmetic and metacognitive monitoring domain-specific? *PLOS ONE*, *15*, e0229932. https://doi.org/10.1371/journal.pone.0229932

Bernard, S., Proust, J., & Clément, F. (2015). Procedural metacognition and false belief understanding in 3-to 5-year-old children. *PloS One*, *10*, e0141321. https://doi.org/10.1371/journal.pone.0141321

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three-and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, *107*, 1018–1034. https://doi.org/10.1016/j.cognition.2007.12.008

Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability through adaptive training. *Journal of Experimental Psychology: General*, *148*, 51–64. https://doi.org/10.1037/xge0000505

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*, 121–138. https://doi.org/10.1017/S0140525X09000545

De Gardelle, V., Le Corre, F., & Mamassian, P. (2016). Confidence as a common currency between vision and audition. *PloS One*, *11*, e0147901. https://doi.org/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. https://doi.org/10.1177/0956797614528956

Dehaene, S. (2011). *The number sense: How the mind creates mathematics, revised and updated edition*. Oxford University Press.

Dehaene, S., Lau, H., & Kouider, S. (2017). What is consciousness, and could machines have it? *Science*, *358*, 486–492. https://doi.org/10.1126/science.aan8871

Einav, S., & Robinson, E. J. (2010). Children's sensitivity to error magnitude when evaluating informants. *Cognitive Development*, *25*, 218–232. https://doi.org/10.1016/j.cogdev.2010.04.002

Fleming, S. M. (2017). HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *1*, nix007. https://doi.org/10.1093/nc/nix007

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, *10*, 843–876. https://doi.org/10.3758/BF03196546

Geurten, M., Meulemans, T., & Lemaire, P. (2018). From domain-specific to domain-general? The developmental path of metacognition for strategy selection. *Cognitive Development*, *48*, 62–81. https://doi.org/10.1016/j.cogdev.2018.08.002

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling uncertain and acting on it during the preschool years: A metacognitive approach. *Child Development Perspectives*, *7*, 160–165. https://doi.org/10.1111/cdep.12035

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, *16*, 1–14. https://doi.org/10.1017/S0140525X00028636

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26–37. https://doi.org/10.2307/1130386

Gopnik, A., & Bonawitz, E. (2015). Bayesian models of child development. *Wiley Interdisciplinary Reviews. Cognitive Science*, *6*, 75–86. https://doi.org/10.1002/wcs.1330

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences*, *113*, 3492–3496. https://doi.org/10.1073/pnas.1515129113

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley & Sons, Inc.

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. https://doi.org/10.1037/a0012682

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. *Evolutionary Origins and Early Development of Number Processing*, 305–333. https://doi.org/10.1016/B978-0-12-420133-0.00012-0

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*, 1768–1776. https://doi.org/10.1177/0956797614542273

Heyes, C., Bang, D., Shea, N., Frith, C. D., & Fleming, S. M. (2020). Knowing Ourselves Together: The Cultural Origins of Metacognition. *Trends in Cognitive Sciences*, *24*, 349–362. https://doi.org/10.1016/j.tics.2020.02.007

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*, 10382–10385. https://doi.org/10.1073/pnas.0812142106

Kim, S., Sodian, B., Paulus, M., Senju, A., Okuno, A., Ueno, M., Itakura, S., & Proust, J. (2020). Metacognition and mindreading in young children: A cross-cultural study. *Consciousness and Cognition*, *85*, 103017. https://doi.org/10.1016/j.concog.2020.103017

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, *76*, 1261–1277. https://doi.org/10.1111/j.1467-8624.2005.00849.x

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*, 609–639. https://doi.org/10.1037/0033-295X.100.4.609

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*, 441–453. https://doi.org/10.1111/j.1467-7687.2009.00907.x

Kuzyk, O., Grossman, S., & Poulin‐Dubois, D. (2020). Knowing who knows: Metacognitive and causal learning abilities guide infants' selective social learning. *Developmental Science*, *23*, e12904. https://doi.org/10.1111/desc.12904

Lecce, S., Demicheli, P., Zocchi, S., & Palladino, P. (2015). The origins of children's metamemory: The role of theory of mind. *Journal of Experimental Child Psychology*, *131*, 56–72. https://doi.org/10.1016/j.jecp.2014.11.005

Lockl, K., & Schneider, W. (2004). The effects of incentives and instructions on children's allocation of study time. *European Journal of Developmental Psychology*, *1*, 153–169. https://doi.org/10.1080/17405620444000085

Lockl, K., & Schneider, W. (2007). Knowledge about the mind: Links between theory of mind and later metamemory. *Child Development*, *78*(1), 148–167. https://doi.org/10.1111/j.1467-8624.2007.00990.x

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*, 1778–1787. https://doi.org/10.1111/j.1467-8624.2011.01649.x

Mamassian, P. (2020). Confidence Forced-Choice and Other Metaperceptual Tasks. *Perception*, *49*, 616–635. https://doi.org/10.1177/0301006620928010

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*, 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Maniscalco, B., & Lau, H. (2014). Signal Detection Theory analysis of Type 1 and Type 2 data: Meta-d', response-specific Meta-d', and the Unequal Variance SDT model. In *The Cognitive Neuroscience of Metacognition* (pp. 25–66). Springer. https://doi.org/10.1007/978-3-642-45190-4_3

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, *88*, 78–92. https://doi.org/10.1016/j.neuron.2015.09.039

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*, 404–418. https://doi.org/10.1037/a0029500

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*, 109–133. https://doi.org/10.1037/0033-2909.95.1.109

Odic, D. (2018). Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Developmental Science*, *21*, e12533. https://doi.org/10.1111/desc.12533

Odic, D., & Starr, A. (2018). An introduction to the Approximate Number System. *Child Development Perspectives*, *12*, 223–229. https://doi.org/10.1111/cdep.12288

O'Leary, A. P., & Sloutsky, V. M. (2017). Carving metacognition at its joints: Protracted development of component processes. *Child Development*, *88*, 1015–1032. https://doi.org/10.1111/cdev.12644

Paulus, M., Tsalas, N., Proust, J., & Sodian, B. (2014). Metacognitive monitoring of oneself and others: Developmental changes during childhood and adolescence. *Journal of Experimental Child Psychology*, *122*, 153–165. https://doi.org/10.1016/j.jecp.2013.12.011

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*, 366–374. https://doi.org/10.1038/nn.4240

Rahnev, D., Desender, K., Lee, A. L. F., Adler, W. T., Aguilar-Lleyda, D., Akdoğan, B., Arbuzova, P., Atlas, L. Y., Balcı, F., Bang, J. W., Bègue, I., Birney, D. P., Brady, T. F., Calder-Travis, J. M., Chetverikov, A., Clark, T. K., Davranche, K., Denison, R. N., Dildine, T., … Zylberberg, A. (2020). The Confidence Database. *Nature Human Behaviour*, *4*, 317–325. https://doi.org/10.1038/s41562-019-0813-1

Rohwer, M., Kloo, D., & Perner, J. (2012). Escape from metaignorance: How children develop an understanding of their own lack of knowledge. *Child Development*, *83*, 1869–1883. https://doi.org/10.1111/j.1467-8624.2012.01830.x

Rouault, M., McWilliams, A., Allen, M. G., & Fleming, S. M. (2018). Human metacognition across domains: Insights from individual differences and neuroimaging. *Personality Neuroscience*, *1*, e17. https://doi.org/10.1017/pen.2018.16

Shea, N., & Frith, C. D. (2019). The global workspace needs metacognition. *Trends in Cognitive Sciences*, *23*, 560–571. https://doi.org/10.1016/j.tics.2019.04.007

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review*, *120*, 779–797. https://doi.org/10.1037/a0034191

Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances*, *2*, 2398212818810591. https://doi.org/10.1177/2398212818810591

van Loon, M., & van de Pol, J. (2019). Judging own and peer performance when using feedback in elementary school. *Learning and Individual Differences*, *74*, 101754. https://doi.org/10.1016/j.lindif.2019.101754

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. https://doi.org/10.1177/0956797614538458

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind tasks. *Child Development*, *75*, 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. https://doi.org/10.1016/0010-0277(83)90004-5