

# Knowledge of development during pregnancy and the transition to parenthood: Psychometric properties of the Domains of Development Instrument

International Journal of  
Behavioral Development  
1–9

© The Author(s) 2024



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/01650254241265586  
journals.sagepub.com/home/jbd



Merideth Gattis<sup>1</sup> , Quoc Cuong Truong<sup>2</sup>, Carol Cornsweet Barber<sup>2</sup>,  
Wendy Middlemiss<sup>3</sup> and Oleg N. Medvedev<sup>2</sup>

## Abstract

Researchers and practitioners need robust measurement tools for evaluating knowledge of child development to better support parents and their children during pregnancy and the transition to parenthood. We addressed this need by evaluating the psychometric properties of the Domains of Development Instrument (DoDI) for measuring knowledge of developmental milestones from birth to 3 years. We evaluated four types of validity evidence for the DoDI: test content, response processes, internal structure, and relations to other variables. We convened an expert panel to evaluate test content and conducted cognitive interviews with mothers to evaluate response processes. We also collected responses from a sample of 418 English-speaking pregnant women to evaluate internal structure and relations to other variables. We observed content validity and response process validity, as well as the predicted internal structure, internal consistency, test–retest reliability, and convergent validity. We conclude with recommendations for future research with the DoDI.

## Keywords

Developmental milestones, knowledge of development, parenting knowledge, pregnancy, reliability, transition to parenthood, validity

One of the most important ways to support families and promote healthy child development is to help parents and other caregivers acquire accurate knowledge about child development (Marope & Kaga, 2015). Our scientific understanding remains limited, however, about fundamental issues such as whether and how caregiver knowledge varies across different aspects, or domains, of development. Accurate assessment of caregiver knowledge is essential for determining first whether and when variations in knowledge predict parenting beliefs and behaviors and related child outcomes, and second, how helping parents gain knowledge can help parents create contexts that facilitate positive child outcomes. To address current limitations, researchers and practitioners need accurate and robust measurement tools for evaluating knowledge across different areas of development (Bartlett et al., 2018; National Academies of Sciences, Engineering, and Medicine, 2016). We addressed the need for a robust tool for assessing knowledge across different domains of development by evaluating the psychometric properties of an existing measure, the Domains of Development Instrument (DoDI, Tamis-LeMonda et al., 2002).

## Knowledge of Development

Parents, policymakers, and researchers consider knowledge of development an important factor in supporting healthy child

development. For example, more than 60% of parents taking part in a national survey in the United States agreed that research about child development can help them be better parents (Zero to Three, 2016). The information that new parents most frequently seek is an understanding of what infants and young children can do and when; a considerable majority of parents report searching for information about developmental milestones, including physical, mental, social, and emotional capabilities (Zero to Three, 2018). Parents' search priorities are consistent with recommendations from policymakers and researchers that knowledge of developmental milestones can help parents anticipate children's needs, improve the detection of developmental delays, and support parent well-being during the transition to parenthood (Bartlett et al., 2018; Benasich & Brooks-Gunn, 1996; National Academies of Sciences, Engineering, and Medicine, 2016; Staal, 2016; Veddovi et al., 2001).

<sup>1</sup>Cardiff University, UK

<sup>2</sup>University of Waikato, New Zealand

<sup>3</sup>University of North Texas, USA

### Corresponding author:

Merideth Gattis, School of Psychology, Cardiff University, Park Place, Cardiff CF10 3AT, UK.

Email: gattism@cardiff.ac.uk

Despite evidence that parents want to have accurate knowledge of child development, there appear to be gaps in parents' understanding of what children can do and when. Many parents have less accurate knowledge about children's capabilities during infancy than childhood; one researcher described this as parents expecting "too little, too late" (Epstein, 1980; Sommer et al., 1993; Zero to Three, 2016). In a U.S. survey, parents expected children to first experience complex emotions at a later age than they do and to benefit from people talking to them or reading to them at later ages than they do (Zero to Three, 2016). Similarly, in a population-based study in Turkey, most mothers overestimated the ages at which children achieve most developmental milestones, including social smiling, vocalizing, and sitting with support (Ertem et al., 2007).

Evidence of knowledge across different aspects of development is variable but broadly indicates that parents know less about some aspects of development compared to others. In one study, adolescent mothers taking part in a family support program in New York City had less accurate knowledge of the social and play milestones that children achieve between birth and 3 years compared to cognitive, language, and motor milestones (Tamis-LeMonda et al., 2002). In another study, adults in Alberta, Canada, had less accurate knowledge of cognitive, emotional, and social development compared to motor development from birth to 6 years (Rikhy et al., 2010). Similarly, Jordanian mothers had less accurate knowledge of cognitive and emotional milestones than physical milestones from birth to 1 year (Safadi et al., 2016). In all three studies, parents demonstrated greater knowledge of motor development than other domains, but the domains for which their knowledge was less accurate varied. Generalizing is difficult because each study used a different measure.

## Current Approaches to Measurement

Researchers have developed numerous measures of knowledge of development, and assessment approaches have varied widely. Some measures use open-ended questions (e.g., "When do children begin to vocalize in response to someone talking to them?"; Ertem et al., 2007). Other measures use multiple choice questions (e.g., "Most infants can walk alone by [a] 2 months, (b) 9 months, (c) 13 months, or (d) 24 months"; Sommer et al., 1993, p. 391) or checklists ("By 3 months of age, most babies can" followed by a list of capabilities and the instruction to tick all that apply; Reich, 2005). The most common format is a statement followed by a rating scale, but rating scales vary, even for use of the same instrument. For example, the Knowledge of Development Inventory (KIDI) includes the statement "Most infants are ready to be toilet trained by one year of age," which some researchers evaluate with the options agree, younger, older, or unsure, and other researchers evaluate with a 5-point Likert-type scale ranging from strongly agree to strongly disagree (Bornstein et al., 2010; Lefever et al., 2008; MacPhee, 2002; Nuttall et al., 2015).

Definitions of knowledge of development differ across measures as well. Most instruments have included not only knowledge of developmental achievements but also more general attitudes and beliefs about children (e.g., "All infants need the same amount of sleep"; Barboza-Salerno, 2020). Several measures have also included beliefs about how to care for children (e.g., "You can spoil babies if you soothe them every time they cry"; Gozali et al., 2020). These wide-ranging operationalizations of

knowledge of development have implications for the design and interpretation of studies investigating associations between knowledge of development and other caregiver characteristics. For example, numerous studies have used the measures listed above to evaluate whether knowledge of development is related to parenting attitudes and beliefs, in which case overlapping measurement would de facto increase the likelihood of observing relations between the purportedly distinct constructs (e.g., Scarzello et al., 2016).

## Improving Measurement

Child Trends, an organization focused on improving child outcomes, has called for researchers and practitioners to use more rigorous methods to evaluate parenting knowledge (Bartlett et al., 2018). Researchers and policy organizations have also called for developmental scientists to increase participant diversity to include people of any gender from a wider range of cultural, ethnic, educational, and economic backgrounds (e.g., Bartlett et al., 2018; Nielsen et al., 2017; Singh et al., 2023). These two aims are intertwined because rigorous methods for evaluating knowledge of development need to be accessible and valid for respondents regardless of their age, educational, or cultural background.

The capacity of assessment instruments to be useful across diverse groups rests in part on the instrument format. Thus, an especially important consideration for measuring knowledge of development is item format. Open-ended questions are easier to understand, especially when used in spoken interviews, and therefore less likely to introduce bias related to literacy or other education issues (Ertem et al., 2007). However, analyzing responses to open-ended questions requires additional processing steps, is less efficient, and can introduce error. In contrast, multiple-choice questions and rating scales enable more efficient and uniform analytic strategies, but the cognitive demands for participants and the potential for sociocultural biases are higher (Schwarz, 2007). One solution is to combine open-ended questions (e.g., "When do children begin to vocalize in response to someone talking to them?"; Ertem et al., 2007) with a continuous response scale. This combined approach can help to reduce bias and error and has the additional benefit that continuous response scales yield interval rather than ordinal data (Rioux & Little, 2020; Truong et al., 2024).

A second consideration for improving the measurement of knowledge of development is psychometric evaluation. Few studies of knowledge of development have reported psychometrics other than internal consistency reliabilities, which vary widely (MacPhee, 2002; Orme & Hamilton, 1987). Studies that have evaluated the factor structure of knowledge measures have identified inconsistencies between design intent and observed structure (e.g., Orme & Hamilton, 1987). The most widely used tool for evaluating knowledge of development, the KIDI, was designed to measure four types of parenting knowledge (developmental milestones, developmental processes, parenting beliefs, and health and safety), but psychometric evaluation does not support a four-factor model (MacPhee, 2002). Bornstein and colleagues (2020) conducted a rigorous evaluation of the KIDI in different societies and found support for a unidimensional, invariant model of a shortened version consisting of 25 of the original KIDI items, including questions about developmental milestones and parenting beliefs.

A third avenue for improving the measurement of knowledge of development is standardized scoring based on psychometric evidence. MacPhee advised researchers to calculate a total KIDI score across all items, consistent with psychometric evidence, but in practice scoring varies widely, with some researchers reporting a single score (though frequently across different items, for example, Nuttall et al., 2015; Rowe et al., 2016), some researchers reporting overall knowledge and knowledge of developmental milestones separately (e.g., Zand et al., 2015), and other researchers reporting scores separated by domains (e.g., cognition, language, and social-emotional skills, McMillin et al., 2015; Safadi et al., 2016) and/or scores for parenting beliefs and health and safety (e.g., Hamzallari et al., 2023). Similarly, Ertem and colleagues designed the Caregiver Knowledge of Child Development Inventory (CKCDI) to measure two factors, developmental milestones and caregiver behaviors that stimulate development, but factor analysis of the CKCDI indicated that the scale was made up of three factors that did not correspond to the design intent (Ertem et al., 2007). Studies using the CKCDI have reported scores based on design intent rather than factor structure (Ertem et al., 2007; Shrestha et al., 2019).

## Our Study

We addressed the need for a robust measurement tool to assess knowledge of development by evaluating an existing measure, the DoDI. Tamis-LeMonda and colleagues (2002) designed the measure for a study in which they asked low-income adolescent mothers from diverse ethnic backgrounds to estimate the ages at which children were first able to perform 52 activities. They chose everyday activities that usually emerge in typically developing children by the age of 3 years. Further, they focused on items reflecting developmental progress common across different cultural settings. The items were drawn from standardized assessment tools such as the Bayley Scales of Infant Development and from a previous study on mothers' knowledge of play and language (Tamis-LeMonda et al., 1998). The items covered five different domains of development: cognitive, language, motor, social, and play. Each activity was described using simple language and mothers were asked to respond with an age in months.

The DoDI has promising features, including a clear focus on knowledge of evidence-based milestones from different domains of development and a simple format with a continuous response scale, but psychometric evaluation is needed. Following the recommendations from the American Psychological Association, National Council on Measurement in Education, and the American Educational Research Association Standards for Educational and Psychological Testing, we evaluated four types of validity evidence: test content, response processes, internal structure, and relations to other variables (American Educational Research Association et al., 2014).

To examine test content, we asked experts to evaluate the adequacy with which the DoDI items represent developmental milestones from birth to 3 years, including milestones in cognitive, language, motor, social, and play development. Response processes were assessed through cognitive interviews with parents, evaluating whether respondents interpreted DoDI items as intended and used relevant knowledge to respond (Willis, 2005). To further evaluate validity evidence, we asked a large sample of pregnant women to complete the DoDI alongside the

KIDI, a widely used measure of knowledge of development. To examine the internal structure of the DoDI, we evaluated factor structure and internal reliability. We also evaluated test-retest reliability. To examine relations between the DoDI and other variables, we compared DoDI scores with KIDI scores to examine convergent validity and also compared DoDI scores for pregnant women with and without previous parenting experience.

## Method

### Participants and Procedures

All recruitment and study procedures were reviewed and approved by the Cardiff University School of Psychology Research Committee (EC.10.11.02.2661G and EC.21.01.12.6260). All participants provided informed consent. A four-step design was used to assess four aspects of validity, as described below.

**Test Content.** Four experts in child development evaluated test content. Panel members had expertise in education, family sciences, human development, and pediatrics. They evaluated whether DoDI items adequately represent developmental milestones in cognitive, language, motor, social, and play development from birth to 3 years. Panel members considered relevance and comprehensiveness for each of the five domains of development. Panel members were asked "Are these items appropriate for measuring knowledge of development from birth to 3 years?" and "Do the items of each domain fit that domain of development?"

**Response Processes.** Mothers ( $N=4$ ;  $M_{age}=34$  years, range 28–43 years) participated in cognitive interviews to evaluate validity based on response processes, as outlined in the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014). Interviews were held in person or online, depending upon the mothers' stated preference. Mothers read each item aloud and described their thought processes as well as responses, following the procedures for cognitive interviews recommended by Willis (2005). This provided information related to whether people interpret each DoDI item as intended and whether people refer to their knowledge of development when deciding on a response.

**Internal Structure and Convergent Validity.** We recruited 418 English-speaking pregnant women via the participant panel provider Prolific (www.prolific.co), who provided a link to study materials on Qualtrics including the DoDI and the KIDI. We excluded 153 additional participants who did not confirm that they were pregnant or did not complete the study. The primary inclusion criterion was being pregnant because pregnancy is an important period for the development of parenting cognitions and because it allowed us to compare responses from women in similar stages of life but with differing levels of parenting experience (i.e., women expecting their first child versus women expecting their second or third child) (Barboza-Salerno, 2020; Mascheroni et al., 2022).

The final sample ranged from 18 to 45 years ( $M_{age}=29.87$  years,  $SD=5.40$ ). Almost 37% were expecting their first child, 40% already had one child, and 23% already had two or more

children. Most participants lived in the United Kingdom, and the sample characteristics were broadly consistent with the UK population, where the average age of mothers giving birth to their first child is 30.9 years, and 81.7% of the population is White. The exception to this was education: mothers with higher levels of education were over-represented in our sample compared to the UK population. Further demographic details are presented in Supplementary Materials (S1).

## Measures

**Domains of Development Instrument (DoDI).** The DoDI (Tamis-LeMonda et al., 2002) evaluates knowledge of child development by asking people to estimate when the average child is first capable of performing each of 52 different actions (see S2 in Supplementary Materials). The items describe empirically based milestones across five domains: cognitive, language, motor, social, and play development. We used a continuous response scale ranging from 0 to 36 months. As in Tamis-LeMonda et al. (2002), we encouraged participants to make their best guess if they were not sure of the correct answer. Correct responses fell within the developmental window for that behavior in healthy, typically developing children (see S2 in Supplementary Materials). We calculated domain scores as the percentage of correct items from the total number of items in a given domain. We also calculated a total DoDI score as the percentage of correct items from the total number of items. To conduct CFAs and obtain test-retest coefficients, we used respondents' exact age estimates (e.g., for the item "Imitates simple actions like clapping and waving" a person might have indicated 7 months as the earliest age when the average child is first capable of performing the action). This approach preserved the advantages of the continuous response scale and maximized the accuracy of the factor analyses.

**Knowledge of Infant Development Inventory (KIDI).** We chose the KIDI to evaluate convergent validity because it is the most widely used measure of knowledge of infant development (e.g., Bornstein et al., 2020; Nuttall et al., 2015; Rowe et al., 2016). The KIDI measures knowledge of developmental milestones, as well as cognitions about developmental processes, parenting, and health and safety (MacPhee, 2002). Based on MacPhee's recommendation, we calculated a total KIDI score by dividing the number of correct responses by the total number of all KIDI items. All KIDI items, response options, and scoring are presented in Supplementary Materials (see S3).

## Analytic Plan

**Data Screening.** Prior to analyses, we screened data from the 418 participants completing the DoDI and KIDI. We evaluated the distribution of our data for normality. Skewness values for all measured variables were within the acceptable range of  $-2$  to  $+2$ , indicating that the data did not deviate significantly from a normal distribution. This assessment supports the appropriateness of parametric tests.

The Kaiser-Meyer-Olkin (KMO) statistic was calculated to determine the adequacy of our sample size for factor analysis. The KMO measure obtained was .942, which is well above the recommended threshold of .6, indicating that the sample size is

sufficient and that the pattern of correlations is relatively compact. Thus, factor analysis is likely to yield distinct and reliable factors. This test was performed to examine the hypothesis that the correlation matrix is an identity matrix, which would indicate that the variables are unrelated and unsuitable for structure detection. The Bartlett's Test of Sphericity yielded an approximate chi-square value of 9505.396 with 1,326 degrees of freedom and was significant ( $p < .001$ ), indicating that the variables are correlated highly enough to provide a reasonable basis for factor analysis. We checked for outliers by examining the standardized scores and leverage values. Cases with standardized scores exceeding  $\pm 3$  or high leverage values were considered outliers and were not available in the dataset. Scatterplots and residuals were inspected to ensure that the assumptions of linearity and homoscedasticity were met. The pattern and amount of missing data were also assessed. Missing values were found to be random and not exceeding 0.1% for any individual variable and were considered as negligible.

**Confirmatory Factor Analyses.** CFA was conducted using *Jamovi* (version 2.3) with integrated structural equation modeling (SEM) module (The Jamovi Project, 2022). Three CFA models were tested: a basic 1-factor model, a five-factor correlated model, and a hierarchical five-factor model with one general second-order factor. The basic 1-factor model was applied to test whether the DoDI captures an overarching knowledge of development factor. The correlated five-factor model examined the five-factor construct of the DoDI as proposed by the scale developers, representing the five domains of cognitive, language, motor, social, and play development. The final hierarchical five-factor model with one general second-order factor examined whether five DoDI domains treated as facets would together represent the overarching factor of general knowledge about child development.

All CFA models were estimated using the Weighted Least Squares Method (WLSMV). The decision to use the WLSMV method for estimating our CFA models is justified based on several considerations arising from our data screening and the nature of our data. First, WLSMV does not assume normally distributed variables. Second, although our KMO measure indicated a high sampling adequacy, which might suggest that our sample size is adequate, WLSMV is known to produce more accurate estimates with smaller sample sizes compared to Maximum Likelihood (ML), which typically requires larger sample sizes for reliable estimates. Third, WLSMV is particularly useful when dealing with complex models. CFA can be intricate, with multiple factors and cross-loadings. WLSMV is better at handling models with many parameters and can provide more reliable standard errors and chi-square statistics, which are crucial for assessing model fit. Given that our Bartlett's Test of Sphericity yielded a significant result, indicating that our variables are related and suitable for factor analysis, WLSMV offers an advantage by providing a more accurate chi-square test of model fit. In addition, WLSMV tends to require less computational time than other methods like ML, especially for large datasets or complex models. This can be an important practical consideration when estimating models.

We used several statistics to examine the acceptable model fit including the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), and chi-square degree of freedom ratio ( $\chi^2/df$ ). CFA



**Table 1.** Fit Indices for CFA Models.

Models / fit indices	CFI	TLI	SRMR	RMSEA (95%CI)	$\chi^2/df$ ( <i>p</i> )
One-factor basic	.944	.941	.097	.053 (.056,.056)	2.081 (<.001)
Five-factor correlated	.954	.952	.091	.048 (.045, 0.51)	1.885 (<.001)
Hierarchical five-factor	.953	.951	.092	.048 (.045, .051)	1.906 (<.001)

Note. *N* = 418. CFI: Comparative Fit Index; TLI: Tucker–Lewis index; SRMR: Standardized Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation; CI: confidence interval.

model fit indices measure how well the hypothesized or originally constructed model fits the observed data (Hu & Bentler, 1999; Kline, 2016). The criteria for the acceptable model fit firstly included both CFI and TLI greater than 0.95 (Hair et al., 2021). CFI compares the fit of the model with a null model, while TLI is less sensitive to sample size (Hu & Bentler, 1999). Acceptable fit also requires RMSEA below 0.06 (Hu & Bentler, 1999), and SRMR less than 0.10 (Byrne, 2016). SRMR is a measure of the average absolute discrepancy between the observed and model-implied covariance matrices, while RMSEA is a measure of the discrepancy per degree of freedom, which takes into account the complexity of the model. Moreover, the chi-square test compares the model-implied covariance matrix with the observed covariance matrix (Steiger, 1990). Generally, a non-significant chi-square value indicates an excellent fit. However, the chi-square test is sensitive to sample size and may not be a reliable indicator of fit for larger samples (i.e., greater than 300 participants) (Brown, 2015; Kline, 2016). Therefore, the ratio of chi-square to degrees of freedom ( $\chi^2/df$ ) is widely used in CFA to overcome dependence from sample size with values less than 5, indicating an acceptable fit (Kline, 2016).

**Convergent Validity.** To examine convergent validity, we computed Pearson correlations between the full DoDI, the five DoDI domains, and the KIDI. We compared DoDI accuracy (i.e., percentage of correct items for the total DoDI and each of the five domains) with KIDI accuracy (percent of correct items shown in S3). We hypothesized that DoDI total accuracy would correlate with KIDI total accuracy. We did not have hypotheses about potential relations between the DoDI domain scores and the KIDI, in part because the content of the KIDI is so wide-ranging, as described in the introduction.

To examine the relationship between knowledge measured by the DoDI and experience, we used analysis of variance (ANOVA) to compare the accuracy of the DoDI and its five domains for pregnant women with and without previous parenting experience.

## Results

### Test Content

All panel members (*N* = 4) reported that the 52 DoDI items were relevant to and adequately representative of the proposed five developmental domains. Panel members used well-established standards of developmental progress (e.g., Centers for Disease Control and Prevention, 2023). Three panel members noted that several items drew on more than one domain, reflecting the interactive nature of development. For example, the item “Reaches

for objects held in front of him or her” involves both cognitive and motor abilities, and the item “Can pick out specific people and objects in photographs” involves both cognitive and social abilities. Panel members suggested simplifying one item, “Finds objects in a “3 card monte game”—or any game where objects are hidden under cups or bowls that are then mixed up,” by omitting the phrase “3 card monte game.”

### Response Processes

All cognitive interview participants indicated DoDI items were understood easily, even when they were not sure of the correct answer. Interviewees consistently referred to their knowledge of cognitive, language, motor, social, and play milestones as they read and responded to items. Like the expert panel, interviewees also noted that several items involved skills across different domains, although they used less formal language to describe these domains. Some parents also indicated that they were not familiar with the phrase “3 card monte.” This item was modified by omitting this phrase.

### Internal Structure

To examine the internal structure of the DoDI, we first evaluated whether DoDI items corresponded to the five developmental domains identified by Tamis-LeMonda and colleagues (2002). We hypothesized that factor analysis would support a five-factor model as well as a single-factor model representing overall knowledge. Table 1 displays the CFA fit indices for all CFA models including 1-factor basic, 5-factor, and hierarchical 5-factor models using respondents’ age estimates for DoDI items. As can be seen, the first model tested, a basic 1-factor model, showed a marginally acceptable fit with the indices of CFI and TLI just under the .95 cut-off point. Both 5-factor models (i.e., correlated five-factor model and hierarchical five-factor model with one general second-order factor) achieved acceptable fit indices. There was no difference between these 5-factor models, supporting the robustness of the measure. Factor loading estimates and path diagrams for these models are shown in Supplementary Materials (S2 and S4). They reveal no differences in factor loading estimates between all CFA models, with all factor loadings ranged from .38 to .68. These CFA results provided support for the five-factor structure of the DoDI as well as a single-factor total score.

We used respondents’ age estimates to investigate the internal reliability of the full DoDI and its domains. The full DoDI showed excellent internal reliability with a Cronbach’s alpha ( $\alpha$ ) of .95 and McDonalds omega ( $\omega$ ) of .95. The DoDI domains also showed fair to good internal consistency: cognitive milestones

**Table 2.** Pearson Correlations Between the Knowledge of Infant Development Inventory (KIDI), the Full Domains of Development Instrument (DoDI), and the Five DoDI Domains ( $n=418$ ).

	KIDI	Full DoDI	Cognitive milestones	Language milestones	Motor milestones	Social milestones
Full DoDI	.43* [.35, .51]					
Cognitive milestones	.20* [.10, .29]	.55* [.48, .61]				
Language milestones	.29* [.20, .38]	.65* [.59, .70]	.23* [.14, .32]			
Motor milestones	.40* [.32, .48]	.72* [.67, .76]	.22* [.13, .31]	.30* [.21, .39]		
Social milestones	.15* [.06, .24]	.47* [.39, .54]	.14* [.04, .23]	.21* [.12, .30]	.19* [.09, .28]	
Play milestones	.22* [.13, .31]	.62* [.55, .67]	.14* [.04, .23]	.23* [.13, .31]	.31* [.22, .39]	.08 [-.01, .18]

Note. Estimations of 95% confidence intervals are based on Fisher's  $r$ -to- $z$  transformation and are reported in brackets. KIDI: Knowledge of Development Inventory.

\* $p < .01$ .

( $\alpha = .80$  and  $\omega = .79$ ), language milestones ( $\alpha = .86$  and  $\omega = .83$ ), motor milestones ( $\alpha = .84$  and  $\omega = .82$ ), social milestones ( $\alpha = .76$  and  $\omega = .76$ ) and play milestones ( $\alpha = .85$  and  $\omega = .83$ ).

### Test–Retest Reliability

We examined the test–retest reliability of the DoDI and its five domains in our sample with 129 participants who completed the DoDI a second time 1 month later. The full DoDI had fair test–retest reliability, with an intraclass correlation coefficient (ICC) of .71 between the two test occasions. Two domains also had fair test–retest reliability: motor milestones had an ICC of .76 and social milestones had an ICC of .77. The remaining domains (cognitive, language, and play milestones) had acceptable test–retest reliability, with ICCs ranging from .62 to .65.

### Convergent Validity

To examine convergent validity, we compared DoDI accuracy (i.e., percentage of correct items for the total DoDI and each of the five domains) with KIDI accuracy (percent of correct items shown in S3). Total DoDI accuracy correlated significantly with the KIDI. DoDI Motor and Language domains correlated more strongly with the KIDI than Play, Cognitive and Social domains, but all correlations were significant (see Table 2). These results support the convergent validity of the total DoDI and its five domains in measuring knowledge of development.

To further examine validity, we used ANOVA to compare the accuracy of the DoDI and its five domains for pregnant women with and without previous parenting experience. We hypothesized that pregnant women expecting their first child would have less accurate knowledge of developmental milestones compared to women who already had children and, therefore, had firsthand experience observing developmental milestones. Accuracy ranged from 31% to 47% across different domains and with different levels of parenting experience, with all groups having more accurate knowledge of motor milestones and less accurate knowledge of social milestones (see S5). Pregnant women expecting their first child estimated social milestones less accurately compared to pregnant women who already had one child ( $p = .02$ ). There were no other differences.

We also evaluated whether the accuracy of the DoDI and its five domains differed for participants with different levels of

education and professional expertise using ANOVA (for education) and independent t-tests (for professional expertise). Across all educational levels and occupations, we again observed that knowledge was more accurate for motor milestones and least accurate for social milestones (see S5). Participants with different levels of education and professional expertise did not differ in accuracy.

### Discussion

Researchers and practitioners need robust measurement tools to assess knowledge of development during infancy and early childhood (Bartlett et al., 2018). To address this need, we evaluated four types of validity evidence for the DoDI (Tamis-LeMonda et al., 2002), including test content, response processes, internal structure, and relations to other variables. We chose the DoDI as a measure of parent knowledge of development because it has a clear focus on evidence-based milestones of child development during the first 3 years, distinguishes between different domains of development, and has a simple format with a continuous response scale. We evaluated the internal structure and convergent validity of the DoDI with a large sample of pregnant women because pregnancy is an important period of preparation for parenting and because existing evidence indicates that maternal cognitions during pregnancy influence parent and child outcomes (Barboza-Salerno, 2020; Mascheroni et al., 2022). Further, we were able to examine the applicability of the instrument for individuals with differing levels of parenting experience by including both pregnant women expecting their first child and pregnant women expecting their second or third child.

The CFA examined the internal structure of the DoDI using three different models: a 1-factor model, a correlated 5-factor model, and a hierarchical 5-factor model with one general second order factor. All CFA models demonstrated goodness of fit. These results provide support for the originally proposed 5-factor structure of the DoDI and additionally for a general factor reflecting knowledge of development across all five domains. The full DoDI and its domains had fair to excellent internal reliability and fair to acceptable test–retest reliability. Positive correlations between the DoDI and the KIDI provided evidence of convergent validity.

Further support for the value of the DoDI as a tool to assess parent knowledge of child development was found in

the examination of test content and response processes for the instrument. A panel of experts in child development identified the items as suitable and appropriate. Mothers' responses during cognitive interviews also confirmed that the items were well understood and elicited appropriate response processes.

In sum, the results of our evaluation support the use of the DoDI to assess general knowledge of developmental milestones from birth to 3 years, as well as to evaluate and compare knowledge of development for specific domains. Our results support scoring for the five domains following the recommendations of Tamis-LeMonda and colleagues (2002) and additionally support scoring for overall knowledge based on all items.

### Strengths

This study found clear support for discrete measurement of the five developmental domains assessed by the DoDI as well as for overall knowledge of development. This is an advantage over other measures, such as the KIDI and the CKCDI, where design intent and factor structure do not clearly align (Ertem et al., 2007; MacPhee, 2002).

An important feature of the DoDI is that it defines knowledge of child development in a clear and specific manner that does not include general attitudes and beliefs about children or how to care for them. The DoDI's milestone-focused approach to measuring knowledge of development makes it an appropriate instrument for evaluating hypotheses about whether knowledge of child development influences parenting attitudes and beliefs because it avoids the problem of overlapping constructs that have characterized some previous studies (e.g., Scarzello et al., 2016). Other researchers have argued, however, that knowledge of development rightly includes the understanding of developmental processes and how parents and other caregivers might provide for or protect a child in an age-appropriate manner, and some evidence indicates that milestone knowledge and stimulation knowledge make distinct contributions to parenting (Ertem et al., 2007; MacPhee, 2002). In addition, focusing on developmental milestones may take away from learning about more sustained, day-to-day behaviors, such as infant sleep and feeding, or health and safety issues, that are also related to parenting behaviors and child outcomes (MacDowall et al., 2017; Middlemiss et al., 2015; Winstanley & Gattis, 2013). Future studies might consider the value of separate, additional measures to evaluate stimulation knowledge and knowledge of day-to-day behaviors.

Another important feature of the DoDI is that it combines simple statements about developmental milestones with a continuous response scale, making it appropriate for participants with diverse cultural and educational backgrounds. The DoDI is a good candidate tool for studying knowledge of development in more diverse samples because the operational definition of knowledge of development is clear and focused, the response scale is simple and continuous, and it has robust psychometric properties. In addition, although children's developmental achievements are influenced by cultural and social factors, the sequence and timing of milestones are nonetheless relatively similar among healthy children from different places and contexts (Villar et al., 2019). This interpretation is supported by the observation that domain differences in knowledge were similar to those reported by Tamis-LeMonda et al. (2002). Future research should explore the psychometric properties of the DoDI

in other populations, including the potential for DoDI versions in other languages.

### Limitations and Future Directions

We expected to observe more knowledge of development among women with relevant experience, including those who already had children and those with relevant professional expertise, but for the most part, we did not. The only significant difference in knowledge across groups was knowledge of social milestones, which differed between mothers expecting their first child and those who already had one child. The absence of a difference might be due to the scoring system of the DoDI, which uses a developmental window to dichotomize responses as correct or incorrect. Although the DoDI demonstrated strong psychometric properties in our study, using a developmental window to dichotomize responses may limit the sensitivity of the DoDI and, thus, the capacity of the instrument to detect differences between different groups of participants. A more continuous approach to scoring would capitalize on the continuous response scale and be more sensitive (Pituch & Stevens, 2015; Rioux & Little, 2020; Tabachnick & Fidell, 2013). Future research should focus on improving the scoring system for the DoDI. Further, the large number of items in the DoDI may lead to participant fatigue and error. Future studies might develop and validate a shortened version of the DoDI to increase the instrument's practical utility.

Finally, longitudinal evidence is needed as knowledge of milestones may vary with infant age (Tamis-LeMonda et al., 1998). Tamis-LeMonda and colleagues (2002) reported that mothers were more likely to demonstrate accurate knowledge of a milestone if their infant was near the age at which that milestone would typically emerge. Longitudinal evidence would also allow researchers to evaluate the temporal reliability and stability of the DoDI.

### Implications and Applications

The results of our study have significant implications for assessing knowledge of child development to support policy and practice. Accurate knowledge of child development can help parents understand and anticipate children's social, emotional, and cognitive needs (Bartlett et al., 2018). Accurate knowledge of development can also improve the early detection of developmental issues and, as a result, facilitate intervention (Staal, 2016). Knowledge of development may also support parent well-being during the transition to parenthood (Veddovi et al., 2001).

Current evidence suggests that parents know less about some stages of development compared to others and know less about some domains of development compared to others, but generalizing is difficult because of variations and limitations in measurement approaches (Rikhy et al., 2010; Safadi et al., 2016). Our study provided psychometric support for the assessment of knowledge of specific domains as well as general knowledge using the DoDI, both of which can inform the development of targeted interventions and further research on the educational needs of parents in specific developmental domains. Researchers, health care providers, and family support workers can confidently use the DoDI to evaluate expectations for development and, where needed, intervene to support optimal outcomes for parents and children.

## Conclusions

Our study evaluated four types of validity evidence for using the DoDI as a tool for assessing knowledge of child development from birth to 3 years. The results supported the content validity, response processes, internal structure, and convergent validity of the DoDI in pregnant women and parents. Further research is needed to develop and evaluate improvements to the DoDI and to explore the reliability and validity of the DoDI in other populations. Our results have important implications for researchers, educators, and health care providers, with the potential to inform targeted interventions to improve outcomes for parents and children.

## Data Availability Statement

Study data are available at [https://osf.io/xds3g?view\\_only=64888d249595430bb6654cc08e70b7f2](https://osf.io/xds3g?view_only=64888d249595430bb6654cc08e70b7f2).

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Cardiff University and University of Waikato Strategic International Partnership Collaborative Seed Fund awarded to Merideth Gattis, Carol Barber, and Oleg Medvedev.

## Ethical Approval

All study procedures were reviewed and approved by the Cardiff University School of Psychology Research Ethics Committee and adhere to the tenets of the Declaration of Helsinki including informed consent.

## Preregistration

This study was preregistered at <https://osf.io/3y6ht>.

## ORCID iD

Merideth Gattis  <https://orcid.org/0000-0002-8665-7577>

## Supplemental Material

Supplemental material for this article is available online.

## References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Barboza-Salerno, G. E. (2020). Cognitive readiness to parent, stability and change in postpartum parenting stress and social-emotional problems in early childhood: A second order growth curve model. *Children and Youth Services Review*, *113*, Article 104958. <https://doi.org/10.1016/j.chilyouth.2020.104958>
- Bartlett, J. D., Guzman, L., & Ramos-Olzagasti, M. A. (2018). *Parenting knowledge among first-time parents of young children: A research-to-practice brief*. Child Trends. [https://www.childtrends.org/wp-content/uploads/2018/07/ParentingKnowledge\\_ChildTrends\\_July2018.pdf](https://www.childtrends.org/wp-content/uploads/2018/07/ParentingKnowledge_ChildTrends_July2018.pdf)
- Benasich, A. A., & Brooks-Gunn, J. (1996). Maternal attitudes and knowledge of child-rearing: Association with family and child outcomes. *Child Development*, *67*, 1186–1205. <https://doi.org/10.2307/1131887>
- Bornstein, M. H., Cote, L. R., Haynes, O. M., Hahn, C.-S., & Park, Y. (2010). Parenting knowledge: Experiential and sociodemographic factors in European American mothers of young children. *Developmental Psychology*, *46*(6), 1677–1693. <https://doi.org/10.1037/a0020677>
- Bornstein, M. H., Yu, J., & Putnick, D. L. (2020). Mothers' parenting knowledge and its sources in five societies: Specificity in and across Argentina, Belgium, Italy, South Korea, and the United States. *International Journal of Behavioral Development*, *44*, 135–145.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge
- Centers for Disease Control and Prevention. (2023). *Centers for disease control and prevention developmental milestones*. National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention. <https://www.cdc.gov/ncbddd/actearly/milestones/index.html>
- Epstein, A. S. (1980). *Assessing the child development information needed by adolescent parents with very young children* (Final report, grant no. 90-C-1341). Department of Health, Education, & Welfare.
- Ertem, I. O., Atay, G., Dogan, D. G., Bayhan, A., Bingoler, B. E., Gok, C. G., Ozbas, S., Haznedaroglu, D., & Isikli, S. (2007). Mothers' knowledge of young child development in a developing country. *Child: Care, Health and Development*, *33*, 728–737.
- Gozali, A., Gibson, S., Lipton, L. R., Pressman, A. W., Hammond, B. S., & Dumitriu, D. (2020). Assessing the effectiveness of a pediatrician-led newborn parenting class on maternal newborn-care knowledge, confidence and anxiety: A quasi-randomized controlled trial. *Early Human Development*, *147*, Article 105082. <https://doi.org/10.1016/j.earlhumdev.2020.105082>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. SAGE.
- Hamzallari, O., Rosinski, L., Petrenko, A., & Bridgett, D. J. (2023). Mothers' emotion regulation and negative affect in infants: The role of self-efficacy and knowledge of parenting practices. *Children*, *10*, Article 85. <https://doi.org/10.3390/children10010085>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.
- The Jamovi Project. (2022). *Jamovi* (Version 2.3). [www.jamovi.org](http://www.jamovi.org)
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Lefever, J. B., Howard, K. S., Lanzi, R. G., Borkowski, J. G., Atwater, J., Guest, K. C., Ramey, S. L., & Hughes, K. (2008). Cell phones and the measurement of child neglect: The validity of the parent-child activities interview. *Child Maltreatment*, *13*, 320–333. <https://doi.org/10.1177/1077559508320680>
- MacDowall, P. S., Galland, B. C., Campbell, A. J., & Elder, D. E. (2017). Parent knowledge of children's sleep: A systematic review. *Sleep Medicine Reviews*, *31*, 39–47. <https://doi.org/10.1016/j.smrv.2016.01.002>



- MacPhee, D. (2002). *Revised manual for the knowledge of infant development inventory* [Unpublished manuscript]. Colorado State University.
- Marope, P. T. M., & Kaga, Y. (Eds.). (2015). *Investing against evidence: The global state of early childhood care and education*. United Nations Educational, Scientific and Cultural Organization.
- Mascheroni, E., Grassi, M., Bonanomi, A., Sperotto, R., Deeg, S., Hung, S., Xia, R., Ionio, C., Au, T. K., & Gattis, M. (2022). The role of experience in parenting beliefs of British and Italian women during pregnancy. *Infant Mental Health Journal, 43*(6), 835–848. <https://doi.org/10.1002/imhj.22014>
- McMillin, S. E., Hall, L., Bultas, M. W., Grafeman, S. E., Wilmott, J., Maxim, R., & Zand, D. H. (2015). Knowledge of child development as a predictor of mother-child play interactions. *Clinical Pediatrics, 54*(11), 1117–1119. <https://doi.org/10.1177/0009922815581763>
- Middlemiss, W., Yaure, R., & Huey, E. L. (2015). Translating research-based knowledge about infant sleep into practice. *Journal of the American Association of Nurse Practitioners, 27*(6), 328–337.
- National Academies of Sciences, Engineering, and Medicine. (2016). *Parenting matters: Supporting parents of children ages 0-8*. The National Academies Press. <https://doi.org/10.17226/21868>
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31–38. <https://doi.org/10.1016/j.jecp.2017.04.017>
- Nuttall, A. K., Valentino, K., Wang, L., Lefever, J. B., & Borkowski, J. G. (2015). Maternal history of parentification and warm responsiveness: The mediating role of knowledge of infant development. *Journal of Family Psychology, 29*(6), 863–872. <https://doi.org/10.1037/fam0000112>
- Orme, J. O., & Hamilton, M. (1987). Measuring parental knowledge of normative child development. *Social Service Review, 61*(4), 655–669. <http://www.jstor.org/stable/30011932>
- Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. Routledge.
- Reich, S. (2005). What do mothers know? Maternal knowledge of child development. *Infant Mental Health Journal, 26*, 143–156.
- Rikhy, S., Tough, S., Trute, B., Benzies, K., Kehler, H., & Johnston, D. W. (2010). Gauging knowledge of developmental milestones among Albertan adults: A cross-sectional survey. *BMC Public Health, 10*(1), 183.
- Rioux, C., & Little, T. D. (2020). Underused methods in developmental science to inform policy and practice. *Child Development Perspectives, 14*, 97–103. <https://doi.org/10.1111/cdep.12364>
- Rowe, M. L., Denmark, N., Harden, B. J., & Stapleton, L. M. (2016). The role of parent education and parenting knowledge in children's language and literacy skills among White, Black, and Latino families. *Infant and Child Development, 25*, 198–220. <https://doi.org/10.1002/icd.1924>
- Safadi, R. R., Ahmad, M., Nassar, O. S., Alashhab, S. A., AbdelKader, R., & Amre, H. M. (2016). Jordanian mothers' knowledge of infants' childrearing and developmental milestones. *International Nursing Review, 63*, 50–59.
- Scarzello, D., Arace, A., & Prino, L. E. (2016). Parental practices of Italian mothers and fathers during early infancy: The role of knowledge about parenting and child development. *Infant Behavior & Development, 44*, 133–143. <https://doi.org/10.1016/j.infbeh.2016.06.006>
- Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology, 21*, 277–287. <https://doi.org/10.1002/acp.1340>
- Shrestha, M., Ulak, M., Strand, T. A., Kvestad, I., & Hysing, M. (2019). How much do Nepalese mothers know about child development? *Early Child Development and Care, 189*(1), 135–142. <https://doi.org/10.1080/03004430.2017.1304391>
- Singh, L., Cristia, A., Karasik, L. B., Rajendra, S. J., & Oakes, L. M. (2023). Diversity and representation in infant research: Barriers and bridges toward a globalized science of infant development. *Infancy, 28*(4), 708–737. <https://doi.org/10.1111/inf.12545>
- Sommer, K., Whitman, T. L., Borkowski, J. G., Schellenbach, C., Maxwell, S., & Keogh, D. (1993). Cognitive readiness and adolescent parenting. *Developmental Psychology, 29*(2), 389–398.
- Staal, I. I. E. (2016). *Early detection of parenting and developmental problems in young children: A structured dialogue with parents*. Utrecht University.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173–180. [https://doi.org/10.1207/s15327906mbr2502\\_4](https://doi.org/10.1207/s15327906mbr2502_4)
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics: Pearson new international edition*. Pearson Higher Ed.
- Tamis-LeMonda, C. S., Shannon, J., & Spellmann, M. (2002). Low-income adolescent mothers' knowledge about domains of child development. *Infant Mental Health Journal, 23*(2), 88–103.
- Tamis-LeMonda, C. S., Chen, L. A., & Bornstein, M. H. (1998). Mothers' knowledge about children's play and language development: Short-term stability and interrelations. *Developmental Psychology, 34*(1), 115–124.
- Truong, Q. C., Gattis, M., Barber, C. C., Middlemiss, W., Au, T., & Medvedev, O. N. (2024). Applying Rasch methodology to examine and enhance precision of the Baby Care Questionnaire. *Journal of Child and Family Studies, 33*(1), 166–178. <https://doi.org/10.1007/s10826-023-02772-0>
- Veddovi, M., Kenny, D. T., Gibson, F., Bowen, J., & Starte, D. (2001). The relationship between depressive symptoms following premature birth, mothers' coping style, and knowledge of infant development. *Journal of Reproductive and Infant Psychology, 19*(4), 313–323. <https://doi.org/10.1080/02646830127205>
- Villar, J., Fernandes, M., Purwar, M., Staines-Ures, E., Di Nicola, P., Ismail, L. C., Ochieng, R., Barros, F., Albernaz, E., Victora, C., Kunnawar, N., Temple, S., Giuliani, F., Sandells, T., Carvalho, M., Ohuma, E., Jaffer, Y., Noble, A., Gravett, M., . . . Kennedy, S. (2019). Neurodevelopmental milestones and associated behaviours are similar among healthy children across diverse geographical locations. *Nature Communications, 10*, Article 511. <https://doi.org/10.1038/s41467-018-07983-4>
- Willis, G. B. (2005). *Cognitive interviewing: a tool for improving questionnaire design*. Sage Publications, Inc.
- Winstanley, A., & Gattis, M. (2013). The baby care questionnaire: A measure of parenting principles and practices during infancy. *Infant Behavior and Development, 36*, 762–775.
- Zand, D. H., Pierce, K. J., Bultas, M. W., McMillin, S. E., Gott, R. M., & Wilmott, J. (2015). Accuracy of knowledge of child development in mothers of children receiving early intervention services. *Journal of Early Intervention, 37*, 226–240.
- Zero to Three. (2016). *Tuning in: Parents of young children speak up about what they think, know and need*. <https://www.zerotothree.org/our-work/national-parent-survey-tuning-in/>
- Zero to Three. (2018). *Millennial connections: Findings from ZERO TO THREE's 2018 parent survey*. <https://www.zerotothree.org/our-work/national-parent-survey-millennial-connections/>