Early Human Development 157 (2021) 105367

Contents lists available at ScienceDirect

Early Human Development

journal homepage: www.elsevier.com/locate/earlhumdev

Psychometric properties of the ASQ–3 in a nationally representative sample of Uruguay

Lucía Alvarez-Nuñez [a], Meliza González [a], Fanny Rudnitzky [b], Alejandro Vásquez-Echeverría [a,*]

[a] Institute of Fundamentals and Methods in Psychology, Faculty of Psychology, University of the Republic, Uruguay
[b] National Directorate of Evaluation and Monitoring, Ministry of Social Development, Uruguay

ARTICLE INFO

Keywords:
Ages & Stages Questionnaires Third Version
Psychometric properties
Nationally-representative sample
Spanish version
Cognitive development
Motor development

ABSTRACT

Background: The Ages & Stages Questionnaires Third Version (ASQ-3) identifies the risk of developmental delay in children aged 2 to 66 months. The ASQ-3 is available in many languages. However, there is little evidence of the psychometric properties of the Spanish version and using nationally-representative samples.
Aims: This study evaluates the reliability and factor solution of the Spanish version of the ASQ-3 (18- to 54-month questionnaires) in a large, representative sample of Uruguayan children. Besides, it explores the association of ASQ-3 scores with sociodemographic characteristics.
Method: Participants were 4016 main caregivers selected randomly across the country who completed the ASQ-3 for their children. All participants responded to the ASQ-3 and a sociodemographic questionnaire within the context of a government-run survey of child development.
Results: Most versions of the ASQ-3 in Spanish have acceptable-to-good psychometric properties, supporting the 5-factor-solution. Personal-Social and, to a lesser extent, Problem-solving scores were the subscales that showed more suboptimal internal consistency coefficients. Scores showed higher ceiling effects than the original US sample but varied across domains, with Gross Motor showing the highest pattern. Sex and socioeconomic status are associated with scores of most age-versions and subscales of the ASQ-3.
Conclusions: In general, results support the reliability and dimensionality of ASQ-3 scores, but psychometric properties varied across age-version and domains. Overall, earlier versions presented less precision, while the Personal-social domain showed reduced reliability in most age-versions.

1. Introduction

Early identification of developmental delays is crucial for infant and child well-being as timely referral to health services reduces problems of greater chronicity in the future. The Ages & Stages Questionnaires Third Version (ASQ-3) is widely used to identify developmental delay risk in children aged 2 to 66 months. It is composed of 21 questionnaires for different age-in-months ranges. Each questionnaire consists of 30 items, organized in five dimensions: Communication, Gross Motor, Fine Motor, Problem-Solving, and Personal-Social. Scores of ASQ-3 subscales are norm-referenced according to a sample of 18.572 completed questionnaires of United States (US) children [1]. The ASQ-3 has become a global screening scale, with adequate validity and reliability, for identifying the need for further pediatric assessment [2,3]. It is currently available in many languages. However, there is little evidence of the validity or reliability of the Spanish version of the ASQ-3, nor is it clear the use of nationally-representative samples to determine its psychometric properties.

Previous international studies with the ASQ-3 have reported adequate internal consistency coefficients [1,4–9]. However, Personal-Social is the subscale with more suboptimal alpha indices [1,5,6]. Personal-Social also shows low item-total correlations in Portuguese, Zulu, and Nyanja and two negative alphas in the Indian version [5–7]. Internal consistency of the translated versions was generally lower than the original version in English [4]. The ASQ-3 presents moderate-to-high correlations between subscales and with the total score [1,6,7].

Abbreviations: ASQ-3, Ages & Stages Questionnaires Third Version; CFA, confirmatory factorial analysis; ENDIS, Health and Child Development National Survey; SES, socioeconomic status.

* Corresponding author at: University of the Republic, Faculty of Psychology, Institute of Methods in Psychology, Dr. Tristán Narvaja 1674, Montevideo 11200, Uruguay.
E-mail address: avasquez@psico.edu.uy (A. Vásquez-Echeverría).

https://doi.org/10.1016/j.earlhumdev.2021.105367
Received 29 January 2021; Received in revised form 17 March 2021; Accepted 26 March 2021
Available online 1 April 2021

To our knowledge, only three studies evaluated the factorial structure of the ASQ-3 using different approaches [5,8,10]. In Brazil, both exploratory and confirmatory factor analyses confirmed the subscales' unidimensionality; except for Personal-Social (for the 10-, 54-, and 60-month versions), Communication and Fine Motor domains presented a two-factor solution [5]. Using a confirmatory factorial analysis (CFA), a five-factor solution was reached [8]. Using Parallel Analyses and Exploratory Structural Equation Modelling, the five-factor solution was supported, but more precision was obtained with versions for older children [10]. In that study, Communication and Gross Motor were the domains with the best item-level fit while, across age versions, Personal-social showed more misfit in factor analyses [10]. Results using the Spanish version showed adequate sensitivity, specificity, and predictive values in some South American samples [11,12], including a study in Argentina that included all age versions [12].

Studies that report ASQ-3 means by questionnaires show that Gross Motor has higher averages [1,5,7]. Girls had higher scores than boys in all the domains, except for Gross Motor [5], but other studies did not find systematic differences by sex [7,13]. Also, lower socioeconomic status (SES) was related to lower scores on most ASQ-3 subscales (except for Gross Motor), and lower family income was associated with a significantly increased probability of developmental delays in all domains [13,14]. ASQ-3 scores are also negatively correlated with socio-emotional development as measured with the ASQ-SE [15].

Although the ASQ-3 is widely used internationally, none of the studies reporting its psychometric properties used nationally-representative samples. Moreover, recent studies highlight the need to further assess the psychometric properties of ASQ-3 [16]. Most of the previous research analyzed only a few age intervals simultaneously [1,3,5,7,12]. In this context, this study explores the reliability and factor structure of the majority of the ASQ-3 age versions (precisely 11 of the 21 versions) in a nationally-representative sample of Uruguayan children. We also wanted to explore the associations of the ASQ-3 with sociodemographic variables such as sex and SES using a governmental-run survey of children's developmental characteristics in Uruguay.

2. Materials and method

2.1. Participants and procedure

This study is based on a Health and Child Development National Survey (ENDIS) sample, an Uruguayan government-run panel survey developed by the National Institute of Statistics. The first cohort included children from 0 to 59 months who had been surveyed twice (first in 2013 and then in 2015). The second cohort included children from 0 to 47 months who had been interviewed once (in 2018). We merged questionnaires from the three data collection points to reach a bigger sample size for each age version. All technical information about sampling can be accessed at ine.gub.uy/endis. As a summary, for the 2013 cohort, all households with children aged 4 or less from the Continuous Household Survey (a governmental survey of living conditions and demography) were invited to participate. For the 2018 cohort, households invited to participate were randomly selected from the Ministry of Public Health's Child Birth Registry.

We counted 4016 respondents to the ASQ-3 questionnaires; 813 children responded in both waves, which we included, as they were not repeated in the same analyses. We excluded 2 to 16 months questionnaires because the sample size was too small to enable statistical inferences ($n \leq 135$). The 60-months questionnaire was excluded because three Fine Motor subscale items were not recorded due to software issues at the National Institute of Statistics. Sociodemographic characteristics of the sample are presented in Table 1.

The Institutional Review Board of the Faculty of Medicine, University of the Republic (Uruguay) approved the research protocol. The main caregiver (95.2% mothers, 2.6% fathers, 2.2% others) completed the ASQ-3 at the children's homes after giving informed consent. University students trained for the task administered the questionnaire. Interviewers asked items orally and recorded responses on a tablet.

Table 1
Age of use and samples' sex distribution and income by ASQ-3 questionnaire.

Age version (month)	Age range	n	% boys	Income (mean)
18	17–18 mo	155	51.6	889.3
20	19–20 mo	156	43.2	843.6
22	21–22 mo	145	48.3	867.1
24	23 mo −25 mo, 15 days	196	50.8	830.8
27	25 mo, 16 days–28 mo, 15 days	245	45.8	843.1
30	28 mo, 16 days–31 mo, 15 days	257	47.9	882.8
33	31 mo, 16 days–34 mo, 15 days	310	51.8	963.8
36	34 mo, 16 days–38 mo	435	51.2	928.4
42	39–44 mo	681	49.0	1006.4
48	45–50 mo	730	51.5	970.2
54	51–56 mo	706	52.3	943.6

Notes. % = percentage; mo = months; income is reported here as US dollars (estimated to values of 2018).

2.2. Measures

2.2.1. Sociodemographic data

Caregivers provided information on the sex and date of birth of each child. Additionally, caregivers indicated household income in Uruguayan pesos (national currency) that was deflated to December 2018 to correct the variations due to inflation.

2.2.2. Ages & Stages Questionnaire – Third Edition

We used the ASQ-3 adapted to Spanish by the original authors. The primary caregiver of the children completes the ASQ-3. Items have three answer options, depending on child behavior development (No = 0, Not yet = 5, Yes = 10). Subscale scores were classified using US norm-reference standardization. Children with scores below 2 SD from the mean were classified at risk and those who present between −2 and −1 SD in the monitoring zone [1]. In this study, we computed a total score by domain and an overall ASQ-3 score by adding scores from all subscales. The range of the subscales is 0–60.

2.3. Analysis plan and data treatment

We used SPSS and MPLUS 8.1 for analyzing the data. SPSS was used to calculate descriptive statistics and correlations; missing values were imputed through expectation maximization. MPLUS was used to perform the CFA and estimate McDonald's omega; both analyses were made with the weighted least squares with mean and variance adjustment (WLSMV) [17]. The estimation of the correlation coefficient, McDonald's omega, and the CFA were performed with the responses to items dichotomized (yes/not yet; sometimes was coded as not yet) to avoid statistical threats produced by infrequent or never-used response categories. The ceiling effect represents the percentage of children with the maximum possible score. For the CFA, values of comparative fit index (CFI) and Tucker-Lewis index (TLI) were considered acceptable near 0.90; the root mean square error of approximation (RMSEA) ≤ 0.06 and standardized root mean square residual (SRMR) ≤ 0.08 [18]. For internal consistency, McDonald's omega values ≥ 0.70 were considered acceptable (between 0.80 and 0.89 good and from 0.90 excellent) [19,20]. For correlational coefficients, a minimum effect size of practical significance was considered $r = 0.20$; 0.50 was considered moderate and 0.80 strong [21].

Scores from domains with four or more missing values were excluded [22,23]. The average of participants' subscale scores excluded using this criterion was 2.4% (ranged from 0.6% to 5.3% across versions and

subscales). The questionnaire with the highest proportion of subscale scores excluded was the 36 months (Problem-Solving 5.3%; Communication 4.8%; Fine Motor 4.6%, Personal-Social and Gross Motor 4.4% each). On average, all the other age versions had less than 4% of subscales scores excluded by this criterion.

We imputed missing values for subscales with three or more responses. Across age versions, the average of subscale scores with imputed missing values was 2.2% (range: 0% to 8.8%), except for Communication of 27 months (that obtained 14% of subscales scores with imputed items due to a computer error when recording Item 1; this only occurred in wave 1). Additionally, and only for the second wave, data processing software codified as 0 (valid response value) the missing values. Considering the missing data of other developmental instruments used in the survey (and not analyzed here), for which 0 was an appropriate value to identify missing values, we hypothesize this not an issue for analyses.

3. Results

In Table 2, we present the mean and SD of the ASQ-3 per questionnaire. Average scores were higher than 50 in all Gross Motor questionnaires, in Communication from 27 months, in Problem-Solving in 42 and 48 months, in Personal-Social from 33 months, including the 22 months. In Fine Motor, all scores were $\leq$50.

As expected, ASQ-3 scores showed a tendency towards ceiling effects (see Fig. 1) but varied the extent across domains. On Gross Motor, it was observed in all questionnaires and in Communication from the 24 months to the 54 months. To a lesser extent, this effect was observed in Problem-Solving and Personal-Social subscales.

For the 18-month case, negative correlations were observed between Item 4 with 1 ($r = -0.34$) and between Item 4 with 2 ($r = -0.06$) of Communication. There was also a low correlation between Item 1 with Items 5 ($r = 0.04$) and 6 ($r = 0.04$) of Personal-Social and a high correlation between Item 1 and 2 ($r = 0.93$) of Gross Motor.

In the 20-month questionnaire, negative correlations were observed between Item 2 and Item 6 of Problem-Solving ($r = -0.13$), Items 1 and 5 of Communication ($r = -0.10$) and between Item 1 and Item 2 ($r = -0.16$), Item 3 and Item 6 of Personal-Social ($r = -0.02$). At 27-month, there was a low correlation between Item 6 with: 1 ($r = 0.15$) and 4 ($r = 0.14$) of Problem-Solving subscale. In this age version, Gross Motor items presented very high correlations (highest $r = 0.84$), preventing model identification in factor analyses. There also were low associations between Item 6 with Item 4 of Problem-Solving ($r = 0.05$) and between Item 1 with Item 2 ($r = 0.08$) and Item 6 ($r = 0.01$) of Communication. In the 30-month measure, there was a negative correlation between Item 1 with Items 2 ($r = -0.14$), 4 ($r = -0.01$), and between Item 3 with Item 5 ($r = -0.02$) of Gross Motor. In the 33-month questionnaire, there were low correlations in Problem-Solving, specifically between Item 2 with 4 ($r = 0.01$). We evaluated each age interval's CFA and internal consistency by deleting the items mentioned above, but no significant improvements were found in relevant indices and coefficients.

Fit indices of CFA models are shown in Table 3. Most versions of ASQ-3 present a good model fit. However, 18-, 22-, and 54-month versions presented CFI and TLI values somewhat below cutoff criteria.

The mean factor loading was $\geq$0.38 for all the models. In Communication, the average of the factor loading ranged between 0.57 and 0.84; on Gross Motor between 0.52 and 0.78; on Fine Motor between 0.61 and 0.86; in Problem-Solving between 0.43 and 0.76 and in Personal-Social between 0.38 and 0.75. In Personal-Social 63% of omegas were suboptimal, in Problem-Solving 36%, in Gross Motor 18%, and in Communication 10% (see results in supplementary material number 1). The Fine Motor subscale identified the highest percentage of the population at risk in the 24-months version (12.8%); the lowest risk was identified in the Personal-Social subscale on the 33-months version (1.9%). Overall, the 24-month version flagged more risk (see supplementary material number 2).

Correlations between dimensions of the ASQ-3 are shown in Table 4. All coefficients were statistically significant and positive (ranging from 0.12 to 0.52), except Gross Motor and Problem-Solving of the 22-months version. For Gross Motor, 11.4% of the associations were of no practical significance ($r < 0.20$), in Communication 9.1%, in Personal-Social 6.8%, in Problem-Solving 4.5%, and Fine Motor 0%. The correlation of domains with the total score in all cases was moderate to high, ranging from $r = 0.46$ to $r = 0.82$.

Table 5 shows mean differences on the ASQ-3 by children's sex. Differences were statistically significant; in older questionnaires, effect sizes were larger. A larger size effect was observed in Fine Motor from the 48-month version ($\approx$ 0.51). There are statistically significant differences in Gross Motor by sex only in 22- and 54-month questionnaires. Girls had higher scores in all the domains, except for Gross Motor.

In general, higher ASQ-3 mean scores were associated with higher households' per capita income. The statistically significant differences were observed in older children's questionnaires in almost all dimensions except Personal-Social (where there are no differences by income except in the 42-months version). All results are presented in supplementary material number 3.

Regarding socioeconomic status, in most cases, the percentage of children at risk or monitoring zone increases in the lower tertile SES groups. As exceptions, Communication at the 20- and 27-month and of Gross Motor at the 27- and 42-month versions flagged more risk in the higher-SES tertile compared to the lower-SES tertile.

Table 2
Descriptive statistics of ASQ-3 by questionnaire.

Age version (month)	C	GM	FM	ProblemS	PersonalS	Total
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
18	39.6 (13.2)	55.0 (8.8)	50.3 (12.2)	44.9 (12.2)	50.9 (10.0)	240.7 (40.6)
20	44.6 (14.5)	54.5 (10.6)	49.0 (12.1)	47.7 (10.6)	49.0 (10.4)	244.4 (44.8)
22	45.0 (13.5)	52.1 (9.6)	45.8 (11.9)	48.5 (11.0)	52.1 (9.4)	243.4 (36.8)
24	49.2 (14.3)	51.6 (12.3)	48.1 (10.9)	47.1 (12.5)	48.0 (11.6)	244.0 (46.3)
27	50.6 (12.1)	52.3 (11.2)	41.3 (14.1)	49.9 (10.5)	46.4 (11.2)	240.5 (44.1)
30	52.7 (11.9)	53.0 (8.6)	42.3 (15.0)	48.5 (11.9)	49.8 (10.3)	246.3 (42.4)
33	53.5 (10.4)	54.9 (8.4)	42.8 (15.8)	49.9 (11.6)	51.7 (9.8)	252.9 (40.7)
36	53.9 (9.0)	55.9 (8.2)	45.1 (15.1)	50.0 (12.4)	51.8 (9.6)	256.7 (41.6)
42	53.1 (10.2)	56.7 (7.6)	44.9 (14.6)	50.9 (12.5)	52.5 (9.4)	258.1 (41.5)
48	55.5 (9.2)	56.1 (7.3)	44.0 (14.5)	52.0 (11.3)	53.1 (9.2)	260.7 (39.3)
54	54.9 (9.2)	56.6 (7.7)	48.1 (13.3)	48.0 (12.0)	53.4 (9.4)	260.9 (37.8)

Notes. C = communication; GM = gross motor; FM = fine motor; ProblemS = problem-solving; PersonalS = personal-social; SD = standard deviation.

4. Discussion

This study aimed to evaluate the psychometric properties of the ASQ-3 in a nationally-representative sample of Uruguayan children and explore the associations of the ASQ-3 scores with sociodemographic characteristics such as income and sex. By doing this, we wanted to determine if the ASQ-3 is an adequate screening instrument for use in the national pediatric context and, further, if risk indicators could be used as epidemiological indexes. Our study is the first to analyze the ASQ-3 with a representative national sample and in which most of the age versions were assessed simultaneously.

Compared with the original study, we found higher means in all

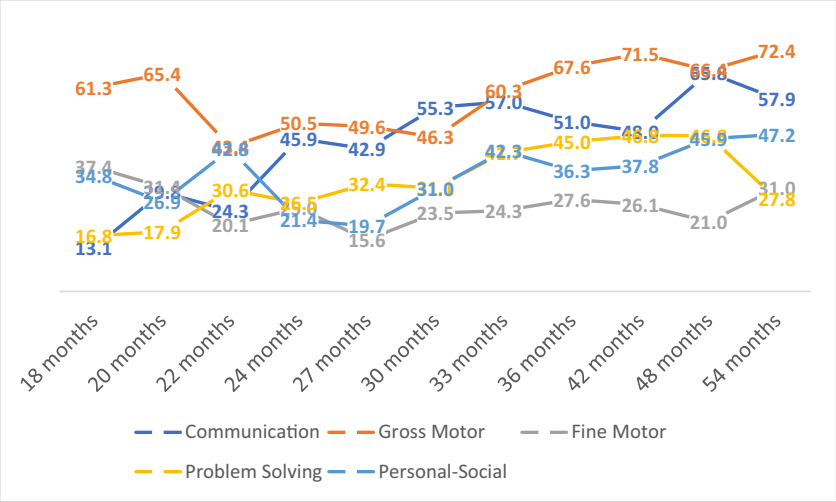

Fig. 1. Ceiling effect by ASQ-3 questionnaire.

Table 3
Model fit of confirmatory factor analysis models.

Age version (month)	χ2	df	CFI	TLI	RMSEA	(IC 90%)		SRMR
18	455.374*	367	0.884	0.871	0.039	0.026	0.051	0.159
20	465.148*	395	0.921	0.913	0.034	0.018	0.046	0.147
22	482.873*	395	0.806	0.786	0.039	0.025	0.051	0.164
24	466.273*	395	0.925	0.917	0.030	0.017	0.041	0.118
27	471.348*	367	0.909	0.899	0.034	0.024	0.043	0.117
30	466.097*	395	0.957	0.953	0.026	0.014	0.036	0.118
33	517.900*	395	0.934	0.928	0.032	0.024	0.039	0.123
36	562.794*	395	0.927	0.920	0.031	0.025	0.037	0.095
42	598.816*	395	0.936	0.929	0.028	0.023	0.032	0.073
48	619.889*	395	0.933	0.926	0.028	0.024	0.032	0.079
54	706.308*	395	0.892	0.882	0.033	0.029	0.037	0.091

Notes. The results of the 18-month questionnaire are without Item 2 in Communication and the 27-month questionnaire are without Item 1 of Gross Motor, because items misfit. Results between the models with all the items and removing misfitting items were quite similar.
* $p < 0.01$.

dimensions and questionnaires (more pronounced in Gross Motor). This pattern also arose in previous studies [1,5,7]. A ceiling effect was also observed on all questionnaires for the Gross Motor subscale, in Communication at 48-month and Personal-Social at 54-month. As expected for a screening tool, these results show that the ASQ-3 provides discriminative scores for children with medium-to-low levels of ability in those domains, but not for higher functioning levels.

Analyses of fit indices and factor loadings support the five-factor solution for all questionnaires as found in previous studies [8,10]. The 22-month questionnaire results should be considered with caution because the sample size is less than 200 [24]. In general, the questionnaires presented acceptable-to-good internal consistency values, as in previous studies [1,4–8]. However, Personal-Social omegas were suboptimal in 63% of the versions. Previous studies also reported a similar trend of psychometric limitations in this subscale [1,5,6,10]. This result suggests that the Personal-Social domain may be more susceptible to cultural differences, such as adults' expectations and understanding in developmental processes. Alternative explanations could be that this domain evaluates a range of varying underlying constructs [7], or that items are more ambiguous or rely more on the observer interpretation of items [10]. For instance, some Personal-Social items refer to adaptation, while others could be considered tapping prosocial behavior. Cultural differences in developmental expectations for adaptive and prosocial behavior and the subscale's content width are both reasons to suggest researchers take with caution Personal-Social scores in Spanish adaptations.

The correlation between domains and total score was positive and statistically significant. These results suggest congruence across the domains assessed, as has been reported previously [1,6]. Gross Motor has the highest percentage of low correlations, probably related with the strong ceiling effect.

Girls obtained higher scores in all ASQ-3 domains except Gross Motor. Gender-related experiences may explain the different cognitive development levels by sex during childhood since adults may relate differently to girls and boys [25]. Concerning SES, ASQ-3 scores were related to income tertiles of the families. Despite that, in some questionnaires, children with lower incomes had higher ASQ-3 scores (e.g. Communication in the 27-month questionnaire). This may be in line with the prevalence of some developmental problems such as autism (that is not related to SES or to ceiling effects that reduce that subscale's discriminant capacity).

In sum, our results suggest that some questionnaires in the ASQ-3 in Spanish provide reliable and valid scores, in terms of dimensionality, for crucial developmental areas, specifically, from 33-month to 54-month age versions. On the contrary, researchers should take with caution scores from versions for ages younger than 33 months. This age effect in measurement accuracy (i.e., older children's questionnaires appear to have better psychometric properties) was previously reported [10]. Similarly, the Personal-Social subscale showed some psychometric limitations since low reliabilities, and item loadings were frequent across versions. As expected, the ASQ-3 has little discriminating capacity for medium and high-developmental levels; and this ceiling effect was mainly observed in the Gross Motor subscale. Therefore, the ASQ-3 scores may help identify risk in children with mild-to-severe

Table 4
Inter-correlations between ASQ-3 subscales and total score.

	18 months					20 months					22 months					24 months				
	2	3	4	5	6.T	2	3	4	5	6.T	2	3	4	5	6.T	2	3	4	5	6.T
1.C	0.22**	0.22**	0.23**	0.14	0.65**	0.33**	0.44**	0.42**	0.34**	0.79**	0.20*	0.22**	0.21*	0.23**	0.69**	0.25**	0.35**	0.34**	0.36**	0.67**
2.GM		0.29**	0.37**	0.35**	0.56**		0.50**	0.42**	0.33**	0.64**		0.37**	0.09	0.19*	0.54**		0.41**	0.28**	0.29**	0.59**
3.FM			0.44**	0.27**	0.64**			0.52**	0.31**	0.74**			0.25**	0.30**	0.69**			0.50**	0.36**	0.73**
4.PS				0.27**	0.69**				0.24**	0.71**				0.22**	0.53**				0.33**	0.72**
5.PersonalS					0.55**					0.60**					0.55**					0.64**

	27 months					30 months					33 months					36 months				
	2	3	4	5	6.T	2	3	4	5	6.T	2	3	4	5	6.T	2	3	4	5	6.T
1.C	0.23**	0.43**	0.39**	0.41**	0.69**	0.18**	0.40**	0.50**	0.27**	0.62**	0.23**	0.44**	0.50**	0.41**	0.70**	0.27**	0.31**	0.44**	0.36**	0.61**
2.GM		0.24**	0.12	0.26**	0.46**		0.35**	0.31**	0.38**	0.56**		0.28**	0.24**	0.31**	0.49**		0.29**	0.31**	0.29**	0.50**
3.FM			0.41**	0.41**	0.79**			0.42**	0.38**	0.81**			0.43**	0.37**	0.79**			0.43**	0.37**	0.78**
4.PS				0.38**	0.67**				0.40**	0.74**				0.43**	0.72**				0.37**	0.76**
5. PersonalS					0.71**					0.64**					0.67**					0.64**

	42 months					48 months					54 months				
	2	3	4	5	6.T	2	3	4	5	6.T	2	3	4	5	6.T
1.C	0.29**	0.42**	0.41**	0.29**	0.65**	0.26**	0.33**	0.39**	0.37**	0.57**	0.19**	0.34**	0.36**	0.18**	0.56**
2.GM		0.28**	0.25**	0.30**	0.46**		0.30**	0.27**	0.31**	0.50**		0.35**	0.27**	0.20**	0.48**
3.FM			0.47**	0.34**	0.82**			0.48**	0.34**	0.82**			0.45**	0.31**	0.78**
4.PS				0.30**	0.73**				0.33**	0.71**				0.28**	0.77**
5.PersonalS					0.58**					0.63**					0.55**

Notes. C = communication; GM = gross motor; FM = fine motor; ProblemS = problem-solving; PersonalS = personal-social; T = total score.
* $p < .05$.
** $p < .01$.

Table 5
Mean scores (and standard deviations) of ASQ-3 subscales by sex.

Age version (month)	Communication			Gross motor			Fine motor		
	Boys	Girls	d	Boys	Girls	d	Boys	Girls	d
18	37.3 (13.4)	41.8 (12.9)	**−0.35**	55.4 (8.1)	54.5 (9.5)	0.10	50.9 (11.4)	50.2 (11.9)	0.06
20	42.3 (15.6)	46.1 (13.6)	−0.26	55.4 (10.9)	53.9 (10.5)	0.14	48.1 (13.3)	49.7 (11.2)	−0.14
22	43.8 (13.9)	46.0 (13.1)	−0.16	53.8 (8.0)	50.6 (10.6)	**0.34**	47.6 (12.1)	44.1 (11.5)	0.30
24	45.7 (15.4)	52.7 (12.3)	**−0.50**	52.8 (9.7)	50.4 (14.5)	0.20	47.6 (10.3)	48.6 (11.4)	−0.09
27	50.2 (12.6)	51.1 (11.8)	−0.07	53.4 (10.6)	51.5 (11.8)	0.17	41.4 (14.1)	41.0 (14.3)	0.03
30	51.4 (13.0)	53.9 (10.8)	−0.22	52.1 (9.4)	53.8 (7.8)	−0.21	39.4 (15.0)	44.9 (14.6)	−0.37
33	52.3 (10.5)	54.8 (10.2)	**−0.25**	54.7 (8.9)	54.9 (8.0)	−0.02	40.1 (16.7)	45.9 (14.1)	**−0.38**
36	52.7 (9.8)	55.1 (8.0)	**−0.27**	55.3 (8.8)	56.4 (7.6)	−0.14	42 (15.9)	48.2 (13.7)	**−0.42**
42	51.7 (11.4)	54.6 (8.7)	**−0.28**	56.5 (8.4)	56.8 (6.7)	−0.04	42.3 (15.6)	47.5 (13.1)	**−0.37**
48	54.3 (10.7)	56.7 (7.1)	**−0.26**	55.6 (8.0)	56.7 (6.4)	−0.15	40.6 (15.2)	47.7 (12.9)	**−0.51**
54	53.7 (10.5)	56.1 (7.3)	**−0.27**	56.0 (8.5)	57.2 (6.7)	**−0.16**	45.5 (14.5)	50.9 (11.1)	**−0.41**

Age version (month)	Problem-solving			Personal-social			Total		
	Boys	Girls	d	Boys	Girls	d	Boys	Girls	D
18	44.7 (12.2)	45.5 (11.9)	−0.07	49.1 (10.9)	52.8 (8.7)	**−0.38**	237.7 (42.3)	244.9 (38.5)	−0.18
20	46.9 (11.4)	48.2 (10.0)	−0.13	45.7 (11.1)	51.4 (9.1)	**−0.57**	237.2 (49.2)	249.5 (40.9)	−0.28
22	50.1 (10.4)	47.1 (11.3)	0.27	51.5 (9.5)	52.7 (9.2)	−0.13	246.5 (34.1)	240.5 (39.0)	0.16
24	46.6 (10.7)	47.7 (14.2)	−0.09	46.6 (10.9)	49.2 (12.1)	−0.23	239.4 (38.2)	248.6 (53.4)	−0.20
27	50.1 (9.9)	49.9 (10.7)	0.02	44.4 (11.1)	47.8 (11.1)	**−0.32**	239.4 (42.4)	241.3 (45.4)	−0.04
30	47.3 (11.9)	49.6 (11.8)	−0.20	47.7 (10.4)	51.9 (9.8)	**−0.41**	238.0 (44.1)	254.1 (39.4)	**−0.39**
33	49.1 (11.7)	50.8 (11.4)	−0.16	49.8 (10)	53.8 (9.2)	**−0.41**	246.2 (41.8)	260.3 (38.2)	**−0.35**
36	48.7 (12.2)	51.3 (12.5)	**−0.21**	49.9 (10.3)	53.7 (8.3)	**−0.40**	248.6 (43.7)	265.1 (37.6)	**−0.40**
42	49.2 (13.6)	52.6 (11.2)	**−0.27**	51.2 (10.5)	53.8 (8.1)	**−0.28**	250.8 (45.8)	265.3 (35.4)	**−0.36**
48	49.8 (12.9)	54.3 (8.9)	**−0.40**	52.0 (9.8)	54.3 (8.3)	**−0.25**	252.3 (43.3)	269.7 (32.5)	**−0.45**
54	46.1 (12.8)	50.0 (10.7)	**−0.33**	52.2 (10.3)	54.8 (8.0)	**−0.28**	253.5 (42.0)	269.0 (30.6)	**−0.42**

Notes. d = Cohen's d; In bold are shown effect sizes that represent statistically significant differences by sex ($p < .05$).

developmental delays, but less valuable to compare developmental levels across groups and across time (e.g. in panel studies), both because of the ceiling effects and different psychometric properties across versions.

Although we analyzed a nationally-representative sample of Uruguayan children, this study presents some limitations. First, the instruments were administered in an interview instead of the self-reported administration suggested in the manual. Second, we worked with a non-clinical sample of children; the instrument may have low sensitivity or discrimination problems in estimating inter-individual differences in behavior problems in the general population. Third, we could not relate ASQ-3 scores with child psychopathology or developmental disorders, being this is the major limitation of our study. Further research trying to validate the Spanish version of the ASQ-3 should take this into account to confirm if the differences found in means scores merit establishing specific cutoff scores based on sensibility and specificity analyses for the Uruguayan population. Finally, we could not examine some age versions of the ASQ-3 due to the reduced sample size.

CRediT authorship contribution statement

Alejandro Vásquez-Echeverría: Conceptualization, Methodology, Supervision, Project administration, Funding acquisition, Writing - review & editing.

Lucia Alvarez-Nuñez: Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Meliza Gonzalez: Data Curation, Software, Formal analysis, Data Curation, Writing – Review & Editing.

Fanny Rudnitzky: Formal analysis, Data Curation, Writing - Review & Editing.

Declaration of competing interest

Fanny Rudnitzky works in the Ministry of Social Development, which runs the governmental survey. The other three authors declare they have no conflict of interest.

Acknowledgments

We thank the Ministry of Social Development of Uruguay for providing us with access to information. This work was funded by an agreement of technical cooperation between the Ministry of Social Development, Uruguay and the Inter-American Development Bank (ATNOC-15774-UR).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.earlhumdev.2021.105367.

References

[1] J. Squires, E. Twombly, D. Bricker, P. LaWanda, Ages & Stages Questionnaires, (ASQ-3): User's Guide, Paul H. Brookes Publishing Co. Inc, Baltimore, MD, 2009.
[2] A. Singh, C.J. Yeh, S. Boone Blanchard, Ages and Stages Questionnaire: a global screening scale, Bol. Med. Hosp. Infant. Mex. 74 (2017) 5–12, https://doi.org/10.1016/j.bmhimx.2016.07.008.
[3] A. Rothstein, A. Miskovic, K. Nitsch, Brief review of psychometric properties and clinical utility of the Ages and Stages Questionnaires, Third Edition for evaluating pediatric development, Arch. Phys. Med. Rehabil. 98 (2017) 809–810, https://doi.org/10.1016/j.apmr.2016.11.001.
[4] T. Velikonja, J. Edbrooke-Childs, A. Calderon, M. Sleed, A. Brown, J. Deighton, The psychometric properties of the Ages & Stages Questionnaires for ages 2–2.5: a systematic review, Child Care Health Dev. 43 (2017) 1–17, https://doi.org/10.1111/cch.12397.
[5] A. Filgueiras, P. Pires, S. Maissonette, J. Landeira-Fernandez, Psychometric properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public child daycare centers, Early Hum. Dev. 89 (2013) 561–576, https://doi.org/10.1016/j.earlhumdev.2013.02.005.
[6] I. Kvestad, S. Taneja, T. Kumar, N. Bhandari, T.A. Strand, M. Hysing, The assessment of developmental status using the Ages and Stages questionnaire-3 in nutritional research in north Indian young children, Nutr. J. 12 (2013) 50, https://doi.org/10.1186/1475-2891-12-50.
[7] C. Hsiao, L. Richter, T. Makusha, B. Matafwali, A. van Heerden, M. Mabaso, Use of the ages and stages questionnaire adapted for South Africa and Zambia, Child Care Health Dev. 43 (2017) 59–66, https://doi.org/10.1111/cch.12413.
[8] S. Ortiz-León, A. Granados-Rojas, J. Cavazos-Olivo, L. De Benito-Avendaño, V. H. Obregón-García, X.M. Duran-Avendaño, R.V. Ramírez-Pérez, J.A. Cárdenas-Medina, B.E. Vargas-Terréz, G. Luna-Guevara, A.C. Rodríguez-Machain, A. L. Jaimes-Medrano, Internal and inter-rater reliability of the ASQ-3 in Mexican

preschoolers, Salud Ment. 41 (2018) 65–72, https://doi.org/10.17711/SM.0185-3325.2018.011.
[9] I. Armijo, L. Schonhaut, M. Cordero, Validation of the Chilean version of the Ages and Stages Questionnaire (ASQ-CL) in community health settings, Early Hum. Dev. 91 (2015) 671–676, https://doi.org/10.1016/j.earlhumdev.2015.10.001.
[10] O.L. Olvera Astivia, B. Forer, G.L. Dueker, C. Cowling, M. Guhn, The Ages and Stages Questionnaire: latent factor structure and growth of latent mean scores over time, Early Hum. Dev. 115 (2017) 99–109, https://doi.org/10.1016/j.earlhumdev.2017.10.002.
[11] L. Schonhaut, I. Armijo, M. Schonstedt, J. Alvarez, M. Cordero, Validity of the ages and stages questionnaires in term and preterm infants, Pediatrics. 131 (2013) e1468–e1474, https://doi.org/10.1542/peds.2012-3313.
[12] A.M.R. Otalvaro, N. Grananã, N. Gaeto, M.D.L.Á. Torres, M.N. Zamblera, M. A. Vasconez, C. Misenta, M.E. Rouviera, J. Squires, ASQ-3: Validation of the Ages and Stages Questionnaire for the detection of neurodevelopmental disorders in Argentine children, Arch. Argent. Pediatr. 116 (2018) 7–13, https://doi.org/10.5546/aap.2018.eng.7.
[13] Q.W. Wei, J.X. Zhang, R.W. Scherpbier, C.X. Zhao, S.S. Luo, X.L. Wang, S.F. Guo, High prevalence of developmental delay among children under three years of age in poverty-stricken areas of China, Public Health 129 (2015) 1610–1617, https://doi.org/10.1016/j.puhe.2015.07.036.
[14] M.R. Potijk, J.M. Kerstjens, A.F. Bos, S.A. Reijneveld, A.F. de Winter, Developmental delay in moderately preterm-born children with low socioeconomic status: risks multiply, J. Pediatr. 163 (2013) 1289–1295, https://doi.org/10.1016/j.jpeds.2013.07.001.
[15] L. Alvarez-Nuñez, M. González, F. Rudnitzky, A. Vásquez-Echeverría, Psychometric properties of the Spanish version of the Ages & Stages Questionnaires: social–emotional in a nationally representative sample, Early Hum. Dev. 149 (2020) 105–157, https://doi.org/10.1016/j.earlhumdev.2020.105157.
[16] J.W. Small, H. Hix-Small, E. Vargas-Baron, K.P. Marks, Comparative use of the Ages and Stages Questionnaires in low- and middle-income countries, Dev. Med. Child Neurol. 61 (2019) 431–443, https://doi.org/10.1111/dmcn.13938.
[17] L.K. Muthén, B.O. Muthén, Mplus User's Guide, 8th ed., Author, Los Angeles, CA, 2017.
[18] L.T. Hu, P.M. Bentler, Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives, Struct. Equ. Model. 6 (1999) 1–55, https://doi.org/10.1080/10705519909540118.
[19] D.V. Cicchetti, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology, Psychol. Assess. 6 (1994) 284–290, https://doi.org/10.1037/1040-3590.6.4.284.
[20] C. Viladrich, A. Angulo-Brunet, E. Doval, Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna, An. Psicol. 33 (2017) 755, https://doi.org/10.6018/analesps.33.3.268401.
[21] C.J. Ferguson, An effect size primer: a guide for clinicians and researchers, Prof. Psychol. Res. Pract. 40 (2009) 532–538, https://doi.org/10.1037/a0015808.
[22] D.B. Rubin, Multiple Imputation for Nonresponse in Surveys Vol. 81, John Wiley, 2004.
[23] J.W. Graham, Missing data analysis: making it work in the real world, Annu. Rev. Psychol. 60 (2009) 549–576, https://doi.org/10.1146/annurev.psych.58.110405.085530.
[24] E.J. Wolf, K.M. Harrington, S.L. Clark, M.W. Miller, Sample size requirements for structural equation models, Educ. Psychol. Meas. 73 (2013) 913–934, https://doi.org/10.1177/0013164413495237.
[25] J. Richter, H. Janson, A validation study of the Norwegian version of the Ages and Stages Questionnaires, Acta Paediatr. 96 (2007) 748–752, https://doi.org/10.1111/j.1651-2227.2007.00246.x.