EUROPEAN JOURNAL OF DEVELOPMENTAL PSYCHOLOGY, 2018
https://doi.org/10.1080/17405629.2018.1480935

DEVELOPMETRICS

OPEN ACCESS Check for updates

Measurement, structural, and functional invariance of parent-child play quality coding across multiple games and parent gender

Bernhard Piskernik and Nina Ruiz

Faculty of Psychology, Department of Applied Psychology: Health: Development, Enhancement and Intervention, University of Vienna, Vienna, Austria

ABSTRACT

Mothers and fathers prefer different games when playing with their children, nonetheless, they share common latent play characteristics. This study aimed to evaluate an observational coding system comprising three rating scales for adjustment, calibration, and familiarity of parent-child play situations that, together, indicate latent parent-child play-quality. 335 toddlers and preschoolers and their parents (332 fathers and 158 mothers) were observed during both physically (father-preferred) and cognitively stimulating (mother-preferred) parent-child games. Measurement, structural, and functional invariance were assessed concerning the type of game and parent gender. No differences were found in any respect. Results showed that the coding system is reliable and unbiased for both parents and game types, indicated equal validity in regard to the external criterion attachment security, and revealed no differences in latent parent-child play-quality across parent gender.

ARTICLE HISTORY Received 22 February 2018; Accepted 20 April 2018

KEYWORDS Measurement invariance; father-mother distinction; parent-child play; observational measure

Parent-child play is one of the major types of interaction between a parent and a child in western cultures (Gauthier & DeGusti, 2012). Thereby it fosters a wide range of child competencies ranging from cognitive to social, emotional, and behavioral skills (Ginsburg, 2007; Lindsey & Mize, 2000; Stgeorge & Freeman, 2017). However, mothers and fathers play differently with their children. While fathers seem to favor physical games to excite, surprise, and encourage their children to take risks (Grossmann et al., 2002; Paquette, 2004), mothers are more

CONTACT Bernhard Piskernik bernhard.piskernik@univie.ac.at
The supplemental data for this article can be accessed at https://doi.org/10.1080/17405629.2018.1480935

likely to structure, guide, and teach in pretend-play situations (Gleason, 2005; John, Halliburton, & Humphrey, 2013).

This seems contradictory to the observed convergence of mothering and fathering towards a general parenting construct in the last decades (Fagan, Day, Lamb, & Cabrera, 2014), but phenotypically different games might share a common latent play quality that is equivalent for both parents. In fact, Ahnert et al. (2017) proposed three observational rating scales of parent-child play as indicators for a latent play quality factor, and demonstrated scalar invariance for fathers and a limited sample of mothers. Although the observed games allowed for pretend play (mother-preferred), they were primarily physical (father-preferred) in nature, therefore it remains unclear whether equivalence of parent-child play quality generalizes to other play types like more cognitively stimulating games. The present study aims to answer this question by evaluating the coding system in terms of its invariance across physical and cognitive games, and both parents. Using confirmatory factor analyses, invariance (see Vandenberg & Lance, 2000) of measurement (configural, metric, scalar, uniqueness, and additionally, due to the dependence of mothers and fathers, covariance of errors) was assessed.

If invariant assessment of parent-child play quality generalizes to other play types, then this would provide a broadly applicable, unbiased measurement instrument of parent-child, which is much needed given the current interest of family research in similarities and differences in parenting between mothers and fathers (see Fagan et al., 2014). Only if measurement invariance holds true, fair group comparisons, which are the base of most research, are possible (Borsboom, 2006). Furthermore, the present study intends to add to the field of comparison of parents by extending the invariance analyses with the structural parameters mean, variance, and covariance to contrast parents regarding level and spread of play quality in different games and its association between parents. At last, functional equivalence regarding the prediction of attachment quality was tested. As a key concept in parent-child interaction (Cassidy & Shaver, 2016), attachment quality is an excellent external validation criterion and the results supplement the ongoing discussion of different formation mechanisms of attachment towards mothers and fathers (e.g., Lucassen et al., 2011).

Methods

Participants

The sample was recruited in Lower Austria and Vienna as part of the CENOF Research Project (Ahnert, Supper, & CENOF, 2014). The present study involved $N = 335$ children (50.9% female) and their parents (332 fathers and 158 mothers). In 154 families both parents at least contributed data on one game. Mothers were undersampled because of the design of the original project, which in its first phase focused solely on fathers and started collecting data on both parents only later on. Since this sampling transition did not change the recruiting strategy of finding parents in Kindergartens and playgroups, we did not expect systematic differences between both sample groups. This assumption is supported by the absence of any significant distributional differences (see online supplement 1). Ages of the target children ranged from 12 to 67 months with m = 29.4 (SD = 16.9) months. Mothers were 35.0 (SD = 5.1) and fathers 38.0 (SD = 6.4) years old on average. The sample represented typical Austrian middle to upper classes, with 53.3% of mothers and 52.7% of fathers holding a bachelor's degree and above. Each household was visited twice. During the visits, a trained research assistant videotaped mothers and fathers while playing the parent-child games and observed them afterwards together with a second research assistant to assess parent-child attachment quality. Mothers and fathers were examined on different days and in randomly assigned orders. All participants gave signed informed consent.

Measures

Play quality

The coding system of parent-child play quality by Ahnert et al. (2017) comprises three 5-point Likert scales assessing: (a) *familiarity* reflecting parent and child feelings of togetherness and joy, and thus the social-emotional quality of the play situation; (b) *calibration* indicating the coordination of the play situation focusing on how play scripts had been mutually developed; (c) *adjustment* capturing structural features of the situation that are seen from the father's and the child's side. For more information on the scales and how they are coded, see online supplement 2. All rated videos were rated by one out of 21 trained research assistants. Interrater reliability was calculated based on 68 father-child and 70 mother-child games and reached an intraclass correlation for the average random rater of $ICC(2,k)$ = .71 for familiarity, .87 for calibration, and .70 for adjustment. In case of multiple ratings the integer value of the median was used in the subsequent analyses.

Games

Play quality was coded based on four different games, one physical and one cognitive game for toddlers (<36 months) and preschoolers (≥36 months), respectively. In the physical game for the younger children, *Candy Bomber* (212 fathers, 115 mothers), parents were asked to hold the child in a vertical position and move him or her around so that a little ball could be picked up and dropped into a bowl. Parents of older children (95 fathers, 31 mothers) were instructed to play *Horse Polo*, which involved carrying the child on their back while the child tried to hit small balls with a long-handled mallet. The cognitive games included *Wild Berry* for the younger children (215 fathers, 105 mothers), where parents hid a small ball (Berry) under one of three cones, like in a classic game of thimblerig. The preschoolers and their parents played *Build-Up* (75 fathers, 29 mothers) and had to build something together out of construction materials. The average game lasted 4 min and 49 s (SD = 2;49). For more information on the games, see online supplement 3.

Attachment

The Attachment Q-Sort (AQS; Waters, 1995) was used to describe the quality of the attachment relationship in a natural setting. The AQS consists of 90 items describing children's attachment behavior in the presence of another person. The items are q-sorted in nine sets ranging from 1 = 'not typical at all' to 9 = 'most typical' of ten items each. The resulting sets were correlated with the set of an ideally securely attached child, providing a score between −1.0 to + 1.0, which was normalized via *r*-to-*z* transformation. Each parent-child relationship was independently rated by two out of 88 trained observers during two-hour home observations. The number of raters is high because the AQS is a key measurement in the original project and several hundred more families were assessed. The interrater reliability in the given data-set was $ICC(2,k) = .95$ and the average score of both raters was used for the subsequent analyses. AQS scores were available for 230 fathers and 124 mothers.

Data analysis

All invariance analyses were conducted within a CFA framework with MPlus 8.0 (Muthén & Muthén, 2017). The CFAs treated item responses as ordered categories and used the robust weighted least squares algorithm WLSMV (Muthén, du Toit, & Spisic, 1997) for model estimation. Since children contributed multiple data points to both levels of comparison, games and parents, all invariance evaluations took the resulting dependence structure into account. The games were modelled as distinct groups with data from the same clusters, while both parents' parent-child play quality was modelled in parallel within each group with correlated errors between the same indicators across parents (see Card, Selig, & Little, 2008).

Model succession

With respect to the two levels of comparison, the sequence of analyses started with a configural model. Assessment of metric (equal factor loadings) and scalar invariance (equal indicator thresholds) across game types and then parents followed, before uniqueness invariance (equal indicator residual variance) for all combinations was evaluated. The last step is not required for unbiased mean value comparisons, but in combination with equal factor variances, uniqueness invariance translates to equal reliability between groups, which we deem a property of interest. Testing measurement invariance was concluded by comparing the indicator error covariances across game types. Next, invariance analyses of the structural parameters factor mean, variance, and covariance followed. Finally, the model was extended by adding the child's attachment security to each parent to the most restrictive accepted model and regressing it on play quality. Invariance of regression paths and error variances was tested across game types and parent gender.

Model evaluation

After controlling for adequate fit of the configural invariance model (CFI ≥ .90 and RMSEA ≤.08; see McDonald & Ho, 2002), testing higher levels of invariance was aimed for. These tests consisted of stepwise fixation of parameter classes across games and parents (see Millsap & Yun-Tein, 2004). The change in model fit was evaluated with χ^2 difference tests using the MPlus function *DIFFTEST* and with the change in CFI (Chen, 2007). Changes in the invariance level with less than or equal to ΔCFI = –.01 and $\alpha < .01$ were considered as acceptable.

Results

The configural invariance model fitted the data excellently without any further adjustments, and none of the consecutive models revealed any indication of non-equivalence of measurement and structural parameters across game types and parents (see Table 1 for all essential results and online supplement 4 for additional fit indices and all parameter estimates). Hence, the coding system of parent-child play quality qualitatively assesses the same construct due to equal factor loadings, without bias due to equal thresholds, and with the same precision due to equal uniqueness and factor variance across cognitive and physical games, and mothers and fathers. Further, the play quality differed neither in level, spread, nor in covariation represented by factor mean, factor variance, and factor covariance, respectively, between parents in any combination. See Figure 1 for the parameters of the final model.

The results of the assessment of functional invariance were strikingly similar. The initial model, which amended the final model of equal factor covariances with AQS scores of both parents, fitted the data well, $\chi^2(79) = 124.0$, $p < .001$, CFI = .990, and RMSEA = .043, and neither fixing all regression paths, $\chi^2(82) = 122.8$, $p = .002$, CFI = .991, RMSEA = .040, and ΔCFI = .001, $\Delta\chi^2(3) = 1.8$, $p = .61$, nor the error variances, $\chi^2(85) = 125.5$, $p = .003$, CFI = .991, RMSEA = .039, and ΔCFI = .000, $\Delta\chi^2(3) = 2.3$, $p = .52$, deteriorated the fit significantly. In the most restrictive model play quality regressed attachment security consistently with $\beta = .30$, $p < .001$. Accordingly, the play quality coding can regress the same percentage of AQS score variance ($R^2 = .09$) for both parents, regardless of the underlying game type. For additional fit indices and all parameter estimates see online supplement 4.

Table 1. Test of measurement and structural invariance

	Model fit					Difference test[a]					
Model	df	χ^2	p	RMSEA	CFI	vs.	Model changes	ΔCFI	$\Delta\chi^2$	Δdf	p
Measurement invariance											
1. configural	10	12.4	.26	.028	.999	–		–	–	–	–
2. metric (games)	14	13.8	.47	.000	1.0	1	Equal factor loadings between games	.001	1.3	4	.86
3. scalar (games)	36	41.0	.26	.021	.999	2	Equal indicator thresholds between games	−.001	27.3	22	.20
4. metric (parents)	38	44.4	.22	.023	.999	3	Equal factor loadings between parents	−.000	3.7	2	.16
5. scalar (parents)	49	57.9	.18	.024	.998	4	Equal indicator thresholds between parents	−.001	13.9	11	.24
6x. residuals free	43	54.0	.12	.029	.998	-	Free indicator residual variance	−.000	–	–	–
6. uniqueness[b]	49	57.9	.18	.024	.998	6x	Equal indicator residual variance	−.000	3.5	6	.74
7. residual covariance	52	60.7	.19	.023	.998	6	Equal indicator residual covariance	−.000	2.9	3	.41
Invariance of structural parameters											
8. mean	55	65.6	.15	.025	.998	7	Equal factor means	−.000	3	4.6	.21
9. variance	58	74.2	.07	.030	.996	8	Equal factor variances	−.002	3	6.1	.11
10. covariance	59	84.1	.02	.037	.994	9	Equal factor covariances	−.002	1	4.2	.04

Notes: [a]Due to the WLMSV estimation a corrected $\Delta\chi^2$ was computed by the MPlus function DIFFTEST.
[b]Equal to model 5, but nested in and compared to a different model (6x not 4).

Discussion

Despite mothers' and fathers' divergent game preferences and play strategies (Gleason, 2005; Grossmann et al., 2002; John et al., 2013; Paquette, 2004), the coding system by Ahnert et al. (2017) unbiasedly captures parent-child play quality for both parents and across game types. Interestingly, play quality differed neither in level nor in relation to the external criterion attachment security between parents. This aligns with the progressing convergence of aspects of mothering and fathering towards parenting in the last decade (Fagan et al., 2014).

Results, however, are limited as they are based on a quite homogeneous sample from primarily middle to upper class from Austria. Thus, future research must focus on measurement invariance for parents with low income, from ethnic minorities, and those who are under-aged or separated. We will also need to investigate non-western cultures with more distinct gender roles in parenting. Further, given the limited sample size of mothers, no invariance analyses were conducted regarding the additional distinction criteria age and gender of the child, which are of interest for future research as well. Apart from that, we do not expect any impairment due to missing cases, because they originated from design rather than sampling bias.

Meanwhile, the coding system of parent-child play quality by Ahnert et al. (2017) can be considered as a reliable, valid, in respect to its association with attachment security, and unbiased measurement for mothers and fathers across different game types.

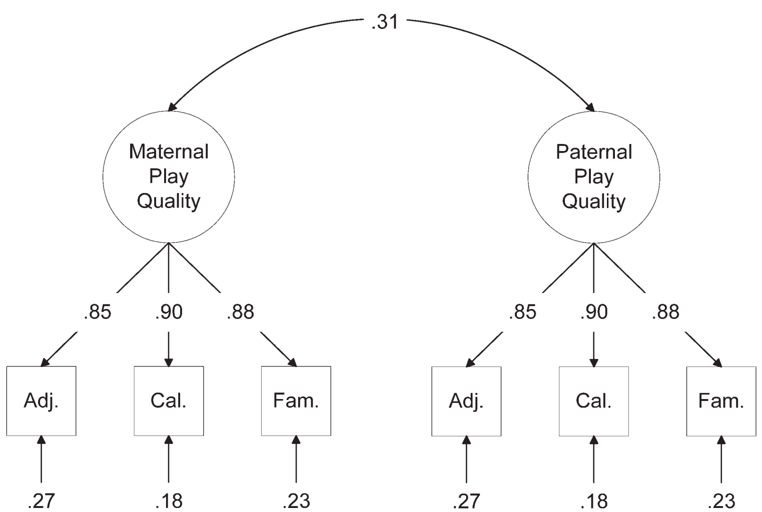

Figure 1. Standardized final model of maternal and paternal parent-child play quality
Note: Correlations between the errors for the same items across parents are not displayed because none was significant.

Author note

This work has been carried out as part of CENOF (the Central European Network on Fatherhood with headquarters at the University of Vienna) and was supported by the Jacobs Foundation (AZ: 2013–1049).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Jakobs Foundation [grant number AZ: 2013-1049].

References

Ahnert, L., & Supper, B., & CENOF. (2014). Father-child attachment as central to the CENOF Research Study. *Attachment and Complex Systems, 1*, 33–42.

Ahnert, L., Teufl, L., Ruiz, N., Piskernik, B., Supper, B., Remiorz, S., … Nowacki, K. (2017). Father-child physical play during the preschool years and child internalizing behaviors: Between robustness and vulnerability. *Infant Mental Health Journal*, 743–756. doi: 10.1002/imhj.21679

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*, 425–440. doi:10.1007/s11336-006-1447-6

Card, N. A., Selig, J. P., & Little, T. D. (2008). *Modeling dyadic and interdependent data in the developmental and behavioral sciences*. New York, NY: Routledge.

Cassidy, J., & Shaver, P. R. (2016). *Handbook of attachment: Theory, research, and clinical applications* (3rd ed.). New York, NY: The Guilford Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi:10.1080/10705510701301834

Fagan, J., Day, R., Lamb, M. E., & Cabrera, N. J. (2014). Should researchers conceptualize differently the dimensions of parenting for fathers and mothers? *Journal of Family Theory & Review, 6*, 390–405. doi:10.1111/jftr.12044

Gauthier, A. H., & DeGusti, B. (2012). The time allocation to children by parents in Europe. *International Sociology, 27*, 827–845. doi:10.1177/0268580912443576

Ginsburg, K. R. (2007). The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics, 119*, 182–191. doi:10.1542/peds.2006-2697

Gleason, T. R. (2005). Mothers' and fathers' attitudes regarding pretend play in the context of imaginary companions and of child gender. *Merrill-Palmer Quarterly, 51*, 412–436. doi:10.1353/mpq.2005.0022

Grossmann, K., Grossmann, K. E., Fremmer-Bombik, E., Kindler, H., Scheuerer-Englisch, H., & Zimmermann, A. P. (2002). The uniqueness of the child-father attachment relationship: Fathers' sensitive and challenging play as a pivotal variable in a 16-year longitudinal ltudy. *Social Development, 11*, 301–337. doi:10.1111/1467-9507.00202

John, A., Halliburton, A., & Humphrey, J. (2013). Child-mother and child-father play interaction patterns with preschoolers. *Early Child Development and Care, 183*, 483–497. doi:10.1080/03004430.2012.711595

Lindsey, E. W., & Mize, J. (2000). Parent-child physical and pretense play: Links to children's social competence. *Merrill-Palmer Quarterly, 46*, 565–591.

Lucassen, N., Tharner, A., Van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Volling, B. L., Verhulst, F. C., … Tiemeier, H. (2011). The association between paternal sensitivity and infant-father attachment security: A meta-analysis of three decades of research. *Journal of Family Psychology, 25*, 986–992. doi:10.1037/a0025855

McDonald, R. P., & Ho, M. H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64–82. doi:10.1037//1082-989X.7.1.64

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*, 479–515. doi:10.1207/S15327906MBR3903_4

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, B., du Toit, S. H., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modelling with categorical and continous outcomes*. Los Angeles, CA: University of California, Los Angeles.

Paquette, D. (2004). Theorizing the father-child relationship: Mechanisms and developmental outcomes. *Human Development, 47*, 193–219. doi:10.1159/000078723

Stgeorge, J., & Freeman, E. (2017). Measurement of father-child rough-and-tumble play and its relations to child behavior. *Infant Mental Health Journal, 38*, 709–725. doi:10.1002/imhj.21676

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4–69. doi:10.1177/109442810031002

Waters, E. (1995). Appendix A: The attachment Q-Set (Version 3.0). *Monographs of the Society for Research in Child Development, 60*, 234–246. doi:10.1111/j.1540-5834.1995.tb00214.x