scientific reports

Check for updates

OPEN

Psychometric properties of the Dutch translation of the Child Self-Report Playfulness questionnaire

Rianne L. van de Weitgraven[1]✉, Brenda R. J. Jansen[3], Els Spaargaren[2], Arne Popma[1], Kim J. Oostrom[1] & Tycho J. Dekkers[4,5,6]

Playfulness represents the disposition to play and has important mental health benefits. Children's playfulness is mainly rated by parents, teachers or trained assessors but playfulness is not always reflected in overt behavior. Fortunately, even young children are able to provide a perspective on their playfulness, as illustrated by research with the Child Self-Report Playfulness questionnaire (CSRP). This study aimed to examine the psychometric properties of the Dutch translation of the CSRP. We included 332 children (*M* age = 5.43 years). Internal consistency of the Dutch version of the CSRP was suboptimal at first administration and acceptable during retest. Test–retest reliability (with an interval of 8 weeks) was adequate. Confirmatory factor analyses demonstrated that a one-factor model fits the CSRP, which supports the idea that playfulness in young children is a unidimensional construct. Scores on the CSRP were not significantly correlated with scores on a parent-rated questionnaire and an observation tool measuring playfulness. Further research to discover how informant and administration discrepancies affect playfulness scores is warranted given this lack of convergence. Children's self-perception of playfulness may complement more traditional measures such as parent reports in future research.

Can you imagine a world without play? Play is a joyous activity that is necessary for optimal child development[1]. It is recognized and promoted as a universal children's right by the United Nations High Commission for Human Rights[2]. Within a loving and safe relationship, children can express exploratory play[3]. By doing so, children engage with the environment, interact with others, and learn[1,4]. This allows higher cognitive skills to develop[5]. Play contributes to the flexibility of the brain and behavior[4]. Through play, children learn to gain control over their emotions and behavior[6]. In this way, play prepares children for focused learning as is expected at school and helps them navigate a complex world[5]. In addition to the development of higher cognitive skills, play and playfulness within a child's early relationships also contribute to socio-emotional development including the development of self-confidence, attachment and resilience[7,8].

Where initial research was mainly focused on play, there is now increased attention for its underlying trait playfulness and the influence of playfulness on human development[9,10]. The study of playfulness is complicated by the fact that different definitions of playfulness are used. Some define playfulness as a state of mind, reflecting the extent to which individuals experience a particular moment in their life as playful[7,11,12]. Others define playfulness as a personality trait, a disposition to play[13]. There are also researchers who take an interactionist approach, in which the playfulness trait interacts with the environment to produce play behavior, the state manifestation of playfulness[14]. The use of these different definitions is a source of concern because playfulness refers to internal psychological qualities (e.g., spontaneity) on some occasions, and to overt behavioral or affective expressions (e.g., laughing) on other occasions. Also, identical items are used to measure playfulness and related constructs like humor, diffusing the boundaries between concepts[14,15]. The development of psychometrically sound measures of playfulness is therefore pivotal.

Playfulness, when defined as a personality trait, is positively associated with cognitive abilities such as divergent thinking, self-estimated intelligence, academic success, and creativity[16–18]. Similar to play, playfulness has several mental health benefits for children and adults. First, playfulness facilitates the use of effective coping strategies in stressful situations[19]. That is, more playful individuals perceive situations as less stressful than less

[1]Department of Child and Adolescent Psychiatry & Psychosocial Care, Amsterdam UMC, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. [2]Department of Paediatric Rehabilitation, Amsterdam UMC, Amsterdam, The Netherlands. [3]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. [4]Levvel, Amsterdam, The Netherlands. [5]Accare Child Study Center, Groningen, The Netherlands. [6]Department of Child and Adolescent Psychiatry, University Medical Center Groningen, Groningen, The Netherlands. ✉email: r.l.vandeweitgraven@amsterdamumc.nl

playful individuals as their confidence in their own abilities to overcome a stressor increases with playfulness. While more playful individuals take an active approach in dealing with stressors, less playful individuals are more likely to engage in self-blame and avoidance. Second, adults with higher levels of playfulness experience more positive emotions and fewer negative emotions. This partially counteracts the negative emotional effects of psychological stress[20]. Third, playfulness enhances adaptability by promoting the ability to deal with unexpected situations, probably due to the tendency of more playful individuals to seek novel situations that allow them to try out newly learned skills[14].

Negative aspects of playfulness have also been reported. In playful adults, prudence may be somewhat lower, which can lead to overstepping boundaries in contact with others[21]. In fact, an impulsive expression of playfulness may be perceived as antagonistic and even destructive[22]. In children, playfulness can manifest itself as clownish behavior, which can have a negative impact on relationships with peers and teachers[23,24].

A large number of self-report measures for playfulness are available for adolescents and adults[25]. In contrast, children's playfulness is mainly rated by parents, teachers or trained assessors, using well-validated measures such as the Children's Playfulness Scale (CPS)[26] and the Test of Playfulness (ToP)[27]. These measures depend on observable behavior. Self-reports of children's playfulness are rare, but have several advantages. Involving children in research, taking into account their views and experiences, respects their right to be heard. As the way a child perceives and looks at something is not visible from the outside, asking them specifically about their perceptions can lead to new insights and the adjustment of expectations[28]. Exclusively having parents, teachers or trained raters as informants may therefore provide a limited view on a child's playfulness. In addition, playfulness may differ between contexts (e.g., at school, at home or with friends) and the child is the only informant who experiences their own playfulness in all these different contexts[29,30].

However, there are also difficulties in using self-report measures with young children. For example, one difficulty with self-report measures in young children is that their responses may be less valid[31], because young children's immature cognitive abilities may make it difficult for them to understand and respond to questions[32]. It is therefore important to develop child-friendly self-report measures that are adapted to their linguistic and cognitive abilities.

There are currently two measures that use young children's self-reports of playfulness: the recently published CPS[33] and the Child Self-Report Playfulness questionnaire (CSRP)[30]. Both are only available in English. In the current study, we introduce a Dutch translation of the CSRP, which can be administered in children from an age of five[30]. The CSRP was for a number of years the only instrument available that allowed children between the ages of five and seven to report on their own playfulness. The CSRP assesses children's self-reported playfulness in an age-appropriate way, using language adapted to young children's language abilities, as children are only asked to point to a puppet that is most like them[30]. Barnett´s definition of playfulness as a predisposition to frame (or reframe) a situation in order to amuse, humor and/or entertain oneself (and possibly others) underpins the CSRP[30,34]. Items are based on the self-reported Adult Playfulness Trait Scale (APTS)[35], the teacher-reported Children's Playfulness Scale[26,34] and the teacher-reported Revised Knox Play Scale[36]. In addition to an overall playfulness factor, these instruments measure different components of playfulness. For example, the APTS measures fun-seeking motivation, uninhibitedness and spontaneity[35], and the CPS measures physical spontaneity, social spontaneity, cognitive spontaneity, manifest joy and sense of humor[26,34]. However, confirmatory factor analysis (CFA) of the selected items resulted in a one factor-model (after removal of two items). The CSRP has good internal consistency and the scores are stable across time (up to one year later)[30].

The aim of this preregistered study was to investigate the psychometric properties of the Dutch translation of the CSRP, with the ultimate goal to enable a multi-informant approach in future studies on playfulness. First, we hypothesized that the items of the CSRP measure the same construct and that the covariance among items is due to a single common factor (i.e., high internal consistency, one-factor structure). An alternative hypothesis, based on the measures underlying the CSRP items, is that playfulness has two components. Three two-factor models seem plausible: (1) a model separating taking initiative (e.g., trying to find a way to make something fun) from reactivity (e.g., joining in when someone else starts something fun)[35]; (2) a model differentiating cognitive (e.g., making up a new game to play) from non-cognitive actions (e.g., singing or dancing)[26,36]; (3) a model differentiating a sense of humor (e.g., I tell jokes) and a more general motivation to seek fun (e.g., looking for fun things to do)[26,35]. Second, we hypothesized scores on the CSRP to be consistent across two administrations (8 weeks apart; i.e., test–retest reliability). Third, we hypothesized that CSRP scores were related to scores on more traditional parent-report and observational measures of children's playfulness (CPS, ToP; i.e., convergent validity).

Method

Participants

We informed local schools and afterschool clubs in the Netherlands about the study. Of the 21 schools and afterschool clubs approached, two did not wish to participate and five did not respond. The remaining 14 schools and afterschool clubs agreed to take part. The schools and afterschool clubs that agreed then sent a letter to the parents of their pupils between the ages of four to eight years old outlining the details of the study and requesting parental permission for their child to take part in the study. Of all parents approached, ten did not agree to their child's participation.

A total of 332 children (M_{age} = 5.42 years, SD = 0.85, 51.50% boys) from schools and afterschool clubs in the Netherlands participated. All children were fluent in Dutch. Schools and afterschool clubs varied in geographical location. The Ethics Review Board of the Faculty of Social and Behavioral Sciences, University of Amsterdam, approved the study (filed as 2020-DP-12792). Table 1 shows the characteristics of the sample. The three subsamples shown in Table 1 are explained in more detail in the Procedure section and differ primarily in terms of the instruments that were administered (CSRP, CPS and ToP).

	Subsample 1 ($N=209$)	Subsample 2 ($N=89$)	Subsample 3 ($N=34$)
Age, *M (SD)*	5.27 (0.75)	5.54 (0.89)	6.00 (0.99)
Sex, % boys	52.60%	48.30%	52.90%
Parental education			
Lower %	NA	0.00%	3.00%
Middle %	NA	9.00%	12.10%
Higher %	NA	43.80%	20.30%
University %	NA	47.20%	54.50%
CSRP *M* (*SD*; range)	7.38 (2.35; 0–11)	7.20 (2.40; 0–11)	6.97 (2.18; 0–11)
CPS, *M* (*SD*; range)	NA	47.72 (8.79; 28–69)	48.71 (9.69; 28–66)
ToP, *M* (*SD*; range)	NA	NA	47.32 (12.13; 25–64)

Table 1. Sample characteristics and descriptive statistics on measures of playfulness. *Note.* Range represents the minimum and maximum score in the sample, M = mean, SD = standard deviation. CPS = Children's Playfulness Scale, CSRP = Child Self-Report Playfulness questionnaire, ToP = Test of Playfulness, NA = Not Applicable.

Materials

Sociodemographic characteristics

Parents, both mothers and fathers, reported on their completed level of education. Parental education scores refer to the average of both parents' education levels. The Dutch Verhage scale, consisting of seven categories, was used to classify the level of education[37]. The seven categories ranged from 1 indicating a maximum of six years of primary education to 7 indicating a master or doctoral university degree. Out of these categories, 1 to 4 were classified as lower education, 5 as middle education, 6 as higher education and 7 as university education.

Child Self-Report Playfulness questionnaire (CSRP)

The CSRP is a self-report measure that takes about five minutes to complete[30]. The 11 items of the CSRP were translated into Dutch and back into English by a professional translator. A videorecorded puppet show was used to deliver the items in an age appropriate manner. In the original CSRP, two puppets were used that looked identical and for which the same voice was used. In the current study, we chose two identical hand puppets in the shape of a dog instead of dolls because they appeal to children and reveal no gender and age. During the puppet show, the two puppets introduced themselves, explained the rules of the task and encouraged the child. The task started with two practice items. In the practice items the two puppets each made an opposite statement in a neutral tone, first about a subject unrelated to playfulness (e.g., "I like pizza" vs. "I don't like pizza") and then about a subject related to playfulness (e.g., "I like to play computer games" vs. "I don't like computer games"). During the remainder of the task the two puppets only made opposite statements about subjects related to playfulness. Playful and non-playful statements were counterbalanced across the two puppets. The child was asked to point to the puppet that was most like him/herself. During the assessment of the puppet show, which was displayed on a computer screen, the experimenter was present to record the answers of the child and explain the procedure or prompt the child when necessary. Scores ranged from 0 to 11, higher scores indicated more playfulness.

Children's Playfulness Scale (CPS)

The Children's Playfulness Scale (CPS) is a parent-rated questionnaire measuring the playfulness disposition, which takes about five minutes to complete[26]. The 23 items of the CPS were translated into Dutch and back to English by a professional translator. The CPS was administered online, and settings made it impossible to skip questions. However, parents could stop while completing the questionnaires and leave any uncompleted items blank. The items (e.g., "The child tells funny stories") were rated on a 5-point Likert scale ranging from 1 ("sounds exactly like the child") to 5 ("doesn't sound like the child at all"). A total score and five subscale scores (Physical Spontaneity, Social Spontaneity, Cognitive Spontaneity, Manifest Joy, and Sense of Humor) were calculated. The Physical Spontaneity and Cognitive Spontaneity subscales each contain five items, and the Social Spontaneity, Manifest Joy and Humor subscales each contain four items. Total scores ranged from 23 to 115. Higher scores indicated less playfulness. The CPS has good internal reliability ($\alpha = 0.79–0.89$) and test–retest reliability ($r = 0.84–0.91$)[13]. In the current sample, the CPS was also internally consistent (total score: $\alpha = 0.81$). The internal consistency of the subscales was acceptable (ranging from $\alpha = 0.65$ to $\alpha = 0.71$). Only the internal consistency of the Cognitive Spontaneity subscale was unsatisfactory ($\alpha = 0.40$).

Test of Playfulness (ToP)

The ToP, version 4, is an observation scale measuring playfulness (defined as the disposition to play) reflecting four elements of play: "internal control", "intrinsic motivation" "freedom to suspend reality" and "framing"[38]. Scoring is based on 15 min of free play[32]. Free play took place in the physical education room of the school of the child. The following age appropriate toys were present: a climbing frame, slide, obstacle course, ball, hand puppets, and Lego. Children were allowed to play with every available toy. A teacher was present or nearby to ensure their safety, but the teacher was not in any way involved in the play.

The ToP consists of 28 items (e.g., "Is actively engaged or supports play of others") that were rated on a 4-point Likert scale assessing either extent (proportion of time), intensity and/or skill of the observed play behavior. Ratings vary from 0 ("rarely or never/not/unskilled") to 3 ("almost always/highly/highly skilled")[39]. Because children did not play in a group, but rather in pairs or triplets, one item could not be rated (e.g., "enters a group"). Sum scores were transformed to Rasch scores, with higher scores indicating more playfulness[39,40]. The ToP has established evidence of construct validity and high inter-rater reliability through a Rasch analysis[38].

One assessor (first author) scored all play recordings. To assess inter-rater reliability, a second assessor (third author) scored a random subset of 17% of the children. Before scoring participants' play, the ratings of both assessors were independently calibrated based on training videotapes. After completing the training, ratings of both assessors fell within the 95% confidence interval on the Rasch Measurement Model (A. Bundy, personal communication, February 28, 2022), indicating reliable scoring.

Procedure

Investigating the factor structure of the CSRP required more participants than other aims of the study. To limit participant strain but still achieve the required sample size for each research question, we used subsamples. All participants of the respective subsamples combined were sufficient to investigate the factor structure of the CSRP.

The first subsample consisted of 209 children who only completed the CSRP. The CSRP was administered by a research assistant at school or at an afterschool club in a separate room, to minimize distraction from the task. Of these children, 24 filled out the CSRP again after 8 weeks ($M_{weeks} = 8.03$, $SD = 0.23$) to determine test–retest reliability. For the first subsample, parents/caretakers provided passive informed consent. At any time during the consent process or during the data collection, children were allowed to stop participating.

The second subsample consisted of 89 children who performed a cognitive test after completion of the CSRP, for purposes beyond the scope of the current study. In addition, the parents/caretakers of the children in this subsample filled out the CPS online. Parents/caretakers of the children in the second subsample provided active informed consent: they were informed about the study and signed informed consent for the child's participation.

The third subsample consisted of 34 children who, next to the CSRP, participated in the ToP. Of these children, 5 filled out the CSRP again after 8 weeks to determine test–retest reliability. Free unstructured play between two or three children was observed and video-recorded by two of the authors and a research assistant. They kept an appropriate distance from the children and did not disturb their play. Children were selected randomly. The CSRP and the ToP were administered at the same day, administration of the CSRP always preceded the ToP. Parents/caretakers of the children provided active informed consent.

Data analysis

As preregistered (https://aspredicted.org/W4Y_8VQ), we investigated validity and reliability of the CSRP using tests of internal consistency, factor structure, test–retest reliability and convergent validity. In addition we investigated the inter-rater reliability of the ToP (not preregistered).

Internal consistency was measured using Cronbach's alpha. Acceptable values of Cronbach's alpha range from 0.70 to 0.95[41]. An exploratory factor analysis was pre-registered to investigate the factor structure. We adjusted the method to CFA as this method is preferred and more common when translating an existing instrument[42]. We tested whether the one-factor model had adequate fit measures and we compared the fit measures to those of various two-factor models. We also compared the factor loadings of the one-factor model with those of the one-factor model by Fink et al.[30]. For the CFA, we used the diagonally weighted least squares estimator (DWLS)[43] as well as its robust variant WLSM, which provides the Satorra Bentler chi-square ($SB\chi^2$)[44]. A model was accepted if the majority of fit measures were acceptable (SRMR values equal to or less than 0.08; RMSEA values equal to or less than 0.06; and CFI and TLI values between 0.90 and 0.95;[45]) and superior to alternative models. CFA was performed using the lavaan package[46] in R[47].

Test–retest reliability was studied by computing the correlation between test (Time 1) and retest (Time 2) scores. Convergent validity was studied by computing the correlations between the CSRP, CPS and ToP scores.

We measured inter-rater reliability of the ToP using Cohen's kappa by case. Kappa ranges [0–0.20], [0.21–0.39], [0.40–0.59], [0.60–0.79], [0.80–0.90], and [0.91–1] are considered none, minimal, weak, moderate, strong, and almost perfect, respectively[48].

As we pre-registered an EFA, we aimed to include 300 participants, as this is classified as good (i.e., all our subsamples combined, $N = 332$)[49,50]. For our current CFA a sample size of more than 200 provides sufficient statistical power[51,52]. For the correlations, we conducted a priori power analyses with G*Power ($\alpha = 0.05$, power = 0.80)[53]. For the correlations related to the convergent validity, we expected medium effects. To detect medium effects, a sample size of 85 was considered sufficient. Note that this size was obtained for the CPS (subsample 2, $N = 89$) but not the ToP (subsample 3, $N = 34$). For the correlation between test and retest of the CSRP, we expected a large effect. We again conducted a priori power analysis with G*Power ($\alpha = 0.05$, power = 0.80)[53]. To detect large effects, a sample size of 21 was considered sufficient, which was indeed obtained ($N = 29$).

As preregistered, the mean score of the completed items of a participant was used when less than 10% of items of a questionnaire was missing. When more than 10% was missing, data of that participant on that questionnaire were discarded. This applied to the CPS data of 6 participants.

Results

Descriptive statistics

Descriptive statistics of the playfulness measures (CSRP, CPS, ToP) are displayed in Table 1. Note that the number of participants differed between measures because the three subsamples completed different sets of measures.

Reliability

The internal consistency of the CSRP was suboptimal during the first survey ($a = 0.64$, $N = 332$), and acceptable during the retest, eight weeks later ($\alpha = 0.74$, $N = 29$). We tested whether this difference could be explained by age differences between the retest subsample and the whole sample, as children who also participated in the retest were older on the first assessment ($M = 5.84$, $SD = 0.80$) than the other children ($M = 5.37$, $SD = 0.84$; $t(207) = 2.83$, $p = 0.005$). However, during the first administration, the internal consistency of the retest subsample was very similar ($\alpha = 0.63$) to that of the whole sample.

We used Spearman's correlation for test–retest reliability because the distribution of the CSRP deviated from normality. CSRP scores were moderately consistent across time ($r_s = 0.46$, $p = 0.013$)[54].

Inter-rater reliability of ToP

A subset (17%, $N = 6$) of ToP cases was coded by two trained raters. Cohen's κ was run on each to determine inter-rater agreement. There was weak agreement on two cases κ = 0.42 (95% CI, 0.21 to 0.63), $p < 0.001$ and κ = 0.53 (95% CI, 0.31 to 0.75), $p < 0.001$, moderate agreement on one case, κ = 0.76 (95% CI, 0.57 to 0.940), $p < 0.001$, good agreement on one case, κ = 0.83 (95% CI, 0.67 to 0.98), $p < 0.001$ and almost perfect agreement on two cases, κ = 0.94 (95% CI, 0.83 to 1.0), $p < 0.001$ and κ = 0.95 (95% CI, 0.86 to 1.0), $p < 0.001$. Disagreement was solved by debate.

Validity

Confirmatory Factor Analysis (CFA)

We specified a one-factor model in which all 11 items were specified to load on a single latent factor. The fit indices suggested acceptable model fit, although SRMR was higher than 0.08 (see Table 2). We contrasted the one-factor model with three two-factor models: model "Initiative", with factors Initiative and Reactivity; model "Cognitive" with factors Cognition and Non-cognition; and model "Humor" with factors Humor and General Fun-seeking. The fit measures of the two-factor models were highly similar to those of the one-factor model: all goodness of fit indices, except SRMR, were in the acceptable range for both the one-factor model and all two-factor models. Although the absolute values of the fit indices were slightly more in favor of the two-factor model 'Humor', the differences between the fit indices of the models were very small. The one-factor model was selected because (1) it is most parsimonious[55], (2) it is consistent with the CSRP's definition of playfulness[30], and (3) the high correlation ($r = 0.80$) between the two factors in the humor model suggests that discriminant validity is low[55]. The one-factor model is presented in Fig. 1.

All items significantly loaded onto the latent factor, with factor loadings ranging from 0.32 to 0.80. Furthermore, the latent variable explained a significant amount of variance in all eleven indicators (R^2 range: 0.10–0.64). Compared to the model fit from the Fink et al.[30] study, the main difference is that in our study all 11 items could be retained. Items 6 (*When someone else starts something fun I always join in*) and 11 (*I do silly things to make other people laugh*), which were removed from the CFA in the Fink et al.[30] study due to relatively low factor loadings, had modest to high factor loadings in the current study (see Fig. 1).

Convergent validity

We conducted correlation analyses to determine whether scores on the CSRP were related to scores on the CPS and ToP. We used Spearman's correlation because the distributions of the CSRP and ToP deviated from normality. There was no significant correlation between the CSRP and the CPS scores, $r_s(114) = -0.15$, $p = 0.103$. This was also the case when investigating the subscales of the CPS: Physical Spontaneity $r_s(118) = -0.04$, $p = 0.703$, Social Spontaneity $r_s(117) = -0.11$, $p = 0.261$, Cognitive Spontaneity $r_s(116) = -0.11$, $p = 0.244$, Manifest Joy $r_s(116) = -0.16$, $p = 0.083$ and Humor $r_s(116) = -0.05$, $p = 0.570$. The correlation between the CSRP and the ToP scores was also not significant, $r_s(31) = -0.04$, $p = 0.841$. Exploratory, we computed the correlation between the CPS and ToP scores and this was also not significant, $r_s(30) = -0.05$, $p = 0.774$.

Discussion

This study investigated the psychometric properties of the Dutch translation of the Child Self-Report Playfulness questionnaire (CSRP)[30]. The internal consistency of the Dutch version of the CSRP was suboptimal during first administration and acceptable during retest. Test–retest reliability, with an interval of 8 weeks, was adequate. A one-factor model fits the Dutch version of the CSRP, supporting the idea that playfulness may be a one-dimensional construct in young children[30]. Relations between scores on the Dutch version of the CSRP, CPS, and ToP do not support convergent validity.

Model	χ^2	SB_{χ}^2	*df*	RMSEA	(90% CI)	SRMR	CFI	TLI
1-factor	75.98**	97.33**	44	0.047	0.028–0.064	0.098	0.935	0.919
Initiative	75.18**	96.97**	43	0.048	0.029–0.065	0.098	0.935	0.917
Cognitive[1]	75.88**	97.27**	43	0.048	0.030–0.066	0.098	0.933	0.915
Humor	70.20**	89.98**	43	0.044	0.024–0.062	0.093	0.945	0.929

Table 2. Fit indices of CFA of CSRP. *Note.* χ^2 = chi-square, SBχ^2 = Satorra and Bentler chi-square, *df* = degrees of freedom, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index. ** $p < 0.01$. [1]Covariance matrix of latent variables is not positive definite.

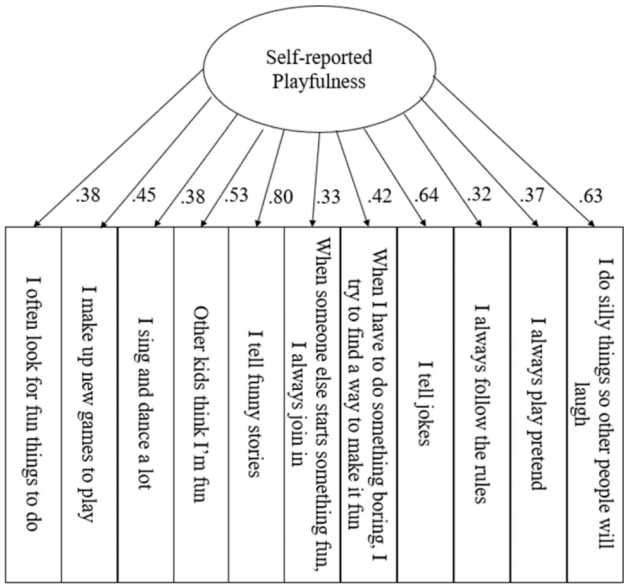

Fig. 1. Standardized parameter estimates from the confirmatory factor analysis with item loadings of the eleven items of the CSRP.

The CSRP has the advantage that young children's covert self-perceptions can be measured. A disadvantage is that the validity of their answers may be suboptimal[30]. Indeed, internal consistency was adequate only at the retest, indicating that children's responses to the CSRP were only moderately consistent. This is in line with previous research on children's self-perception using a puppet interview[56].

The observed increase in internal consistency was unexpected. Possibly, the subsample participating in both test and retest differed from the remaining children. However, at the first administration internal consistency of the whole sample and that of the subsample who also took part in the retest were very similar. Knowing that the stability of personality traits increases substantially from the toddler to the preschool years[57], the increase in internal consistency could be due to the fact that the children were about eight weeks older at the time of the retest. However, the children who took part in the retest were five months older than the other children at the time of the first administration, which makes it seem unlikely that it was just the eight weeks in between that caused the increase in internal consistency, whereas a five-month age difference resulted in very similar internal consistencies. Another possible explanation is that it may have been easier for the children at the retest to understand the questions and link them to personal experiences because they had heard them before, making it easier to answer the questions[58]. Taken together, this may have led to more consistent responses at the retest[59]. However, it is always possible that the difference is due to another cause or random variation.

A one-factor model was found to fit the data, consistent with the findings of Fink and colleagues[30]. However, it is noteworthy that all items were retained in the current study, whereas Fink and colleagues[30] removed two items to obtain a fitting model. It remains unclear how this difference arose, as our study was the closest possible replication of the original study: The sample was approximately the same age and the same procedure was used (except for using puppets instead of dolls).

Unexpectedly, we found no strong relationships between CSRP and other measures of playfulness.The lack of relationships between CSRP scores and scores on other playfulness instruments may complicate the interpretation of what the CSRP measures. This lack of convergent validity may be due to the many differences between the instruments: differences in the underlying definitions of playfulness, methods and informants.

The ultimate goal of this study was to enable a multi-informant approach for future studies on playfulness. Is the CSRP really useful for follow-up research? On the one hand, the CSRP lacks convergent validity. On the other hand, the unidimensionality of the questionnaire, the support for internal consistency at retest and test–retest reliability are important conditions for further research into the validity of the CSRP. In addition, all instruments that measure playfulness have their advantages and disadvantages. For example, good internal consistency is a strength of the parent-reported CPS [34], but the CPS has the disadvantage that parent reports can be biased. For example, parents appeared to rate their child's playfulness higher than a trained assessor[60]. To avoid parental bias, observation measures such as the ToP can be used[38]. The ToP, however, has the disadvantages that the behavior being measured depends on the environment in which the ToP is administered[61] and that the instrument is labor-intensive. Using various instruments simultaneously will most likely provide the most comprehensive understanding of playfulness.

In addition to being part of a multi-informant approach, a psychometrically sound self-report measure such as the CSRP can be used to explore the relationship between children's self-perceptions of their playfulness and other constructs such as cognition. Previous research on children's playfulness as observed by parents and/or teachers has shown that individuals who are more playful are more likely to try out newly learned skills by seeking out novel situations[14]. This raises the question whether playfulness facilitates the acquisition of knowledge and

consequently leads to higher scores on tasks measuring crystallized intelligence. We also know that play enables the development of higher cognitive functions and that playfulness is associated with divergent thinking[5,16]. A relevant direction for future research would therefore be to investigate the link between playfulness and various cognitive skills.

Strengths of the current study are a rigorous and pre-registered approach to the measurement of the psychometric properties of the Dutch version of the CSRP, a large sample, and the inclusion of an observation scale in addition to a parent-rated questionnaire. The use of hand puppets and a dichotomous response format optimized the likelihood of obtaining valid and reliable responses from young children. In fact, there are a number of advantages to using puppets in scientific research with young children. First, the use of puppets helps young children to keep their attention[62,63]. Second, rather than having a conversation with an adult, where young children may be inclined to give the answer they think the other person wants to hear, they may respond more naturally to puppets[56]. Third, it allows shy children to express themselves[56]. In addition, the dichotomous response format used in the CSRP is consistent with young children's dichotomous thinking[64]. The dichotomous response format also avoids the risk of social desirability and the tendency of young children to give extreme responses when presented with more response options, as in a Likert scale[65,66].

Despite these strengths this study also has some limitations. First, the sample included a relatively small number of parents with lower educational levels, which may limit generalizability. Although it is not expected that the underrepresentation of parents with lower education levels had consequences for the psychometric properties of the CSRP as measured, expanding the sample in future research will increase generalizability of the findings. Second, only 34children participated in the play observation (ToP) and there may have been insufficient power to reliably estimate the relationship between the CSRP and ToP. Given the very low correlation, however, it is unlikely that the results would have been very different with a larger sample. Third, the current sample size did not allow for testing whether children of different ages and gender had comparable odds of endorsing the CSRP items or that differential item functioning (DIF) occurred. DIF is a threat to the validity of measures[67]. To avoid jeopardizing the validity of the CSRP, it is recommended to test for DIF in a new sample, as was done by Fink and colleagues (2020)[30]. Fourth, the use of hand puppets has as a possible disadvantage that it may increase task complexity, as seems to be the case when questions are paired with pictures[68]. Fifth, the use of a two-stage response format consisting of two dichotomous responses, resulting in a 4-point response scale, may have resulted in a greater range of responses[33,69].

In conclusion, the current study shows that children's self-perception of playfulness provides a unique perspective. Children's self-perception of playfulness complements more traditional measures such as parent reports and observations. As playfulness is associated with mental health and well-being[21] it is important to assess playfulness early in development. Using the CSRP can aid this assessment.

Data availability

The datasets generated and/or analysed during the current study are not publicly available as not all participants of this study provided written consent for their data to be made publicly available, but are available from the corresponding author on reasonable request.

Received: 12 February 2024; Accepted: 29 October 2024
Published online: 05 November 2024

References

1. Ginsburg, K. R., Committee on Psychosocial Aspects of Child and Family Health. The importance of play in promoting healthy child development and maintaining strong parent-child bonds. *Pediatrics* **119**(1), 182–191 (2007).
2. United Nations General Assembly. Convention on the Rights of the Child. *UN Treaty Ser.* **1577**(3), 1–23 (1989).
3. Fredrickson, B. L. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *Am. Psychol.* **56**(3), 218 (2001).
4. Gordon, G. Well played: The origins and future of playfulness. *Am. J. Play* **6**(2), 234–266 (2014).
5. Vygotsky, L. S. Play and its role in the mental development of the child. *Int. Res. Early Child. Educ.* **7**(2), 3–25 (2016).
6. Berk, L. W., Mann, T. D., & Ogan, A. T. Make-believe play: Wellspring for development of self-regulation. In *Play = Learning: How Play Motivates and Enhances Children's Cognitive and Social-emotional Growth* (eds. Singer, D.G., Golinkoff, R.M. & Hirsch-Pasek, K.) 74–100 (Oxford University Press, 2006).
7. Youell, B. The importance of play and playfulness. *Eur. J. Psychother. Counsell.* **10**(2), 121–129 (2008).
8. Whitebread, D., Coltman, P., Jameson, H. & Lander, R. Play, cognition and self-regulation: What exactly are children learning when they learn through play?. *Educ. Child Psychol.* **26**(2), 40 (2009).
9. Akhtar, S. Normal and Pathological Playfulness. In *Play and Playfulness: Developmental, Cultural, and Clinical Aspects (eds. Akhtar, M.C.)* 69–84 (Rowman and Littlefield, 2011).
10. Proyer, R. T., Brauer, K. & Wolf, A. Assessing other-directed, lighthearted, intellectual, and whimsical playfulness in adults: Development and initial validation of the OLIW-S using self- and peer-ratings. *Eur. J. Psychol. Assess.* **36**(4), 624–634 (2020).
11. Apter, M. *A structural phenomenology of play: A reversal theory approach* (Swets & Beitlinger, 1990).
12. Stenros, J. *Playfulness, play, and games: A constructionist ludology approach* (Tampere University Press, 2015).
13. Proyer, R. T. & Brauer, K. Assessment of playfulness: Current challenges and overview. In *Psychological assessment—Science and practice: Handbook of positive psychology assessment* (eds Ruch, W. et al.) 145–161 (Hogrefe, 2023).
14. Shen, X., Chick, G. & Pitas, N. A. From playful parents to adaptable children: a structural equation model of the relationships between playfulness and adaptability among young adults and their parents. *Int. J. Play* **6**(3), 244–254 (2017).
15. Proyer, R. T. Playfulness and humor in psychology: An overview and update. *Humor* **31**(2), 259–271 (2018).
16. Barnett, L. A. & Kleiber, D. A. Concomitants of playfulness in early childhood: Cognitive abilities and gender. *J. Genet. Psychol.* **141**(1), 115–127 (1982).
17. Bateson, P. & Nettle, D. Playfulness, ideas, and creativity: A survey. *Creat. Res. J.* **26**(2), 219–222 (2014).
18. Proyer, R. T. Being playful and smart? The relations of adult playfulness with psychometric and self-estimated intelligence and academic performance. *Learn. Individ. Differ.* **21**(4), 463–467 (2011).

19. Magnuson, C. D. & Barnett, L. A. The playful advantage: How playfulness enhances coping with stress. *Leisure Sci.* **35**(2), 129–144 (2013).
20. Chang, C. P., Hsu, C. T. & Chen, I. J. The relationship between the playfulness climate in the classroom and student creativity. *Qual. Quant.* **47**(3), 1493–1510 (2013).
21. Proyer, R. T. Playfulness over the lifespan and its relation to happiness. *Z. Gerontol. Geriatr.* **47**(6), 508–512 (2014).
22. Proyer, R. T. & Tandler, N. An update on the study of playfulness in adolescents: its relationship with academic performance, well-being, anxiety, and roles in bullying-type-situations. *Soc. Psychol. Educ.* **23**(1), 73–99 (2020).
23. Barnett, L. A. The education of playful boys: Class clowns in the classroom. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2018.00232 (2018).
24. Fung, W. K. & Chung, K. K. H. Interrelationships among parental play belief, household playfulness, school play behaviors, and social competence of kindergarten children. *J. Leis. Res.* https://doi.org/10.1080/00222216.2023.2297251 (2024).
25. Proyer, R. T. & Jehle, N. The basic components of adult playfulness and their relation with personality: The hierarchical factor structure of seventeen instruments. *Personality Individ. Differ.* **55**(7), 811–816 (2013).
26. Barnett, L. A. The playful child: Measurement of a disposition to play. *Play Cult.* **4**(1), 51–74 (1991).
27. Bundy, A. C. Assessment of play and leisure: Delineation of the problem. *Am. J. Occup. Ther.* **47**(3), 217–222 (1993).
28. Pascal, C. & Bertram, T. Listening to young citizens: The struggle to make real a participatory paradigm in research with young children. *Eur. Early Child. Educ. Res. J.* **17**, 249–262 (2009).
29. De Los Reyes, A. et al. The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol. Bull.* **141**(4), 858 (2015).
30. Fink, E., Mareva, S. & Gibson, J. L. Dispositional playfulness in young children: A cross-sectional and longitudinal examination of the psychometric properties of a new child self-reported playfulness scale and associations with social behavior. *Infant Child Dev.* **29**(4), e2181 (2020).
31. Conijn, J. M., Smits, N. & Hartman, E. E. Determining at what age children provide sound self-reports: An illustration of the validity-index approach. *Assessment.* **27**, 1604–1618 (2020).
32. Eiser, C., Mohay, H. & Morse, R. The measurement of quality of life in young children. *Child Care Health Dev.* **26**, 401–414 (2000).
33. Duss, I., Rüdisüli, C., Wustmann Seiler, C. & Lannen, P. Playfulness from children's perspectives: development and validation of the Children's Playfulness Scale as a self-report instrument for children from 3 years of age. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2023.1287274 (2023).
34. Barnett, L. A. Playfulness: Definition, design, and measurement. *Play Cult.* **3**, 319–336 (1990).
35. Shen, X. S., Chick, G. & Zinn, H. Validating the Adult Playfulness Trait Scale (APTS): An Examination of personality, behavior, attitude, and perception in the nomological network of playfulness. *Am. J. Play* **6**(3), 345–369 (2014).
36. Knox, S. Development and current use of the Knox preschool play scale. In *Play in Occupational Therapy for Children* (eds Parham, L. D. & Fazio, L.) 35–51 (Mosby, 1997).
37. Verhage, F. *Intelligentie en leeftijd bij volwassenen en bejaarden* (Van Gorcum, 1964).
38. Bundy, A. C., Nelson, L., Metzger, M. & Bingaman, K. Validity and reliability of a test of playfulness. *Occup. Ther. J. Res.* **21**(4), 276–292 (2001).
39. Bundy, A. C. *Test of Playfulness, Version 4* (University of Sydney, 2003).
40. Linacre, J. M. *Many-Facet Rasch measurement* (MESA Press, 1989).
41. Tavakol, M. & Dennick, R. Making sense of Cronbach's alpha. *Int. J. Med. Educ.* **2**, 53 (2011).
42. Fokkema, M. & Greiff, S. How performing PCA and CFA on the same data equals trouble. *Eur. J. Psychol. Assess.* **33**(6), 399–402 (2017).
43. Li, C. H. The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychol. Methods* **21**(3), 369 (2016).
44. Satorra, A. & Bentler, P. M. Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika* **75**(2), 243–248 (2010).
45. Hu, L. T. & Bentler, P. M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Modeling* **6**(1), 1–55 (1999).
46. Rosseel, Y. lavaan: An R package for structural equation modeling. *J. Stat. Softw.* **48**(2), 1–36 (2012).
47. R Core Team R: *A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2013).
48. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem. Med.* **22**(3), 276–282 (2012).
49. Comrey, A. L. & Lee, H. B. *A first course in factor analysis* (Psychology Press, 2013).
50. Tabachnick, B. G., Fidell, L. S. & Ullman, J. B. *Using multivariate statistics* (Pearson, 2007).
51. Hoe, S. L. Issues and procedures in adopting structural equation modelling technique. *J. Quant. Methods* **3**(1), 76 (2008).
52. Singh, K., Junnarkar, M. & Kaur, J. *Measures of positive psychology: Development and validation* (Springer, 2016).
53. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**(2), 175–191 (2007).
54. Dancey, C. P. & Reidy, J. *Statistics without maths for psychology* (Pearson, 2007).
55. Brown, T. A. & Moore, M. T. *Handbook of structural equation modeling* (The Guilford Press, 2023).
56. Ringoot, A. P. et al. Young children's self-reported emotional, behavioral, and peer problems: The Berkeley Puppet Interview. *Psychol. Assess.* **25**(4), 1273 (2013).
57. Roberts, B. W. & DelVecchio, W. F. The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychol. Bull.* **126**, 3 (2000).
58. Polit, D. F. Getting serious about test–retest reliability: a critique of retest research and some recommendations. *Qual. Life Res.* **23**, 1713–1720 (2014).
59. McCrae, R. R., Kurtz, J. E., Yamagata, S. & Terracciano, A. Internal consistency, retest reliability, and their implications for personality scale validity. *Pers. Soc. Psychol. Rev.* **15**, 28–50 (2011).
60. Muys, V., Rodger, S. & Bundy, A. C. Assessment of playfulness in children with autistic disorder: A comparison of the Children's Playfulness Scale and the Test of Playfulness. *OTJR Occup. Particip. Health* **26**(4), 159–170 (2006).
61. Bronson, M. R. & Bundy, A. C. A correlational study of a test of playfulness and a test of environmental supportiveness for play. *Occup. Ther. J. Res.* **21**(4), 241–259 (2001).
62. Eder, R. A. Uncovering young children's psychological selves: Individual and developmental differences. *Child Dev.* **61**, 849–863 (1990).
63. Measelle, J. R., Ablow, J. C., Cowan, P. A. & Cowan, C. P. Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Dev.* **69**, 1556–1576 (1998).
64. Gelman, R. & Baillargeon, R. A review of some Piagetian concepts. In *Handbook of child psychology* Vol. 3 (eds Flavell, J. H. & Markman, E. M.) 167–230 (Wiley, New York, 1983).
65. Chambers, C. T. & Johnston, C. Developmental differences in children's use of rating scales. *J. Pediatr. Psychol.* **27**, 27–36 (2002).
66. Kreitchmann, R. S., Abad, F. J., Ponsoda, V., Nieto, M. D. & Morillo, D. Controlling for response biases in self-report scales: Forced-choice vs. psychometric modeling of Likert items. *Front. Psychol.* **10**, 2309 (2019).
67. Chen, W. H., & Revicki, D. Differential item functioning (DIF). *Encyclopedia of Quality of Life and Well-being Research,* 1611–4. (Springer Netherlands, 2014).

68. Marsh, H. W., Ellis, L. A. & Craven, R. G. How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Dev. Psychol.* **38**, 376 (2002).
69. Harter, S. & Pike, R. The pictorial scale of perceived competence and social acceptance for young children. *Child Dev.* **55**, 1969–1982 (1984).

Acknowledgements
We would like to thank all participating schools and after-school clubs for their cooperation. We thank Maria Masseurs, Ilse de Jong, Tynke Kootstra, Aybuke Özbakir and Bas van Tol for helping to make this research happen.

Author contributions
RLvdW, BRJJ, AP, KJO and TJD designed the study. RLvdW coordinated the data-collection. RLvdW analyzed the data, supervised by BRJJ. RLvdW, BRJJ and TJD wrote the main text of the manuscript and ES scored a percentage of the test of playfulness. All authors reviewed the manuscript.

Declarations

Competing interests
The authors report there are no competing interests to declare.

Ethical approval
All experimental protocols were approved by the Ethics Review Board of the Faculty of Social and Behavioral Sciences, University of Amsterdam (filed as 2020-DP-12792).

Informed consent
Parents/caregivers of children in the subsample who only took part in the CSRP provided passive consent. Parents/caregivers of children in the subsample who took part in the CSRP as well as the cognitive tasks and/or the ToP provided active consent. At any time during the consent process or during the data collection, children were allowed to stop participating.

In accordance statement
All methods were carried out by relevant guidelines and regulations and in accordance with the approved ethics application.

Additional information
Correspondence and requests for materials should be addressed to R.L.W.