European Journal of Developmental Psychology

ISSN: 1740-5629 (Print) 1740-5610 (Online) Journal homepage: https://www.tandfonline.com/loi/pedp20

Parent-report measure of Theory of Mind: Adaptation to Polish of the Children's Social Understanding Scale

Marta Białecka-Pikul & Małgorzata Stępień-Nycz

To cite this article: Marta Białecka-Pikul & Małgorzata Stępień-Nycz (2019) Parent-report measure of Theory of Mind: Adaptation to Polish of the Children's Social Understanding Scale, European Journal of Developmental Psychology, 16:3, 318-326, DOI: 10.1080/17405629.2017.1406850

To link to this article: https://doi.org/10.1080/17405629.2017.1406850

EUROPEAN JOURNAL OF DEVELOPMENTAL PSYCHOLOGY
2019, VOL. 16, NO. 3, 318–326
https://doi.org/10.1080/17405629.2017.1406850

Routledge
Taylor & Francis Group

DEVELOPMETRICS

Parent-report measure of Theory of Mind: Adaptation to Polish of the Children's Social Understanding Scale

Marta Białecka-Pikul and Małgorzata Stępień-Nycz

Institute of Psychology, Jagiellonian University, Krakow, Poland

ABSTRACT

The main aim of the study was to investigate whether the Children's Social Understanding Scale (CSUS), a parent report technique, is a valid and reliable measure of Theory of Mind (ToM) abilities in Polish-speaking children. Additionally, the role of language abilities in ToM development was explored. A group of 225 parents of approximately 3.5-year-old Polish children was tested using the Polish version of the CSUS. Their children were tested with a word comprehension test and five behavioural ToM tasks. Satisfactory psychometric properties of the Polish CSUS were observed and positive correlations between the CSUS and behavioural ToM tasks were found. A two-factor structure was found in the CSUS: Mental State Talk (MST) and Mental State Comprehension. The MST factor was significantly related to word comprehension. The utility of the Polish version of the CSUS for future cross-cultural research with this population is discussed.

ARTICLE HISTORY Received 5 May 2017; Accepted 12 November 2017

KEY WORDS Theory of Mind; parent report; validity; reliability; language abilities

The challenge presented by studying individual differences in Theory of Mind (ToM) development has been widely recognized (Hughes & Devine, 2015). The difficulties involved in undertaking such research were partially associated with the dearth of valid, reliable and widely-used scales or other techniques designed for use with children (for exceptions, see e.g.,: Shahaeian, Peterson, Slaughter, & Wellman, 2011). This research gap was recently supplemented by the Children's Social Understanding Scale (CSUS) (Tahiroglu et al., 2014), a parent-report technique measuring ToM abilities in 3-to-6-year-old English-speaking children. However, the results of three studies using this tool reported by Tahiroglu et al. (2014) have 2 important limitations: (1) only English-speaking children were tested; (2) language skills, which are so important in ToM abilities (see: Milligan, Astington, & Dack, 2007), were not taken into account. Our main aim was to adapt the CSUS to the Polish language and find if the CSUS is a valid and reliable

CONTACT Marta Białecka-Pikul marta.bialecka-pikul@uj.edu.pl

Supplemental data for this article can be accessed here [https://doi.org/10.1080/17405629.2017.1406850].

measure outside the context of Western culture. The second aim concerned the relation of language skills to children's ToM abilities, as measured with the CSUS in approximately 3.5-year-old Polish-speaking children.

Tahiroglu et al. (2014) emphasized that the idea of 'singularity' in ToM measurement dominated most research before their study. The authors emphasized that it is not wise 'to focus on a single informant (child) in a single context (laboratory), and even sometimes on a single task type (the false belief task)' (p. 2486). When attempting to achieve convergent validity for a particular measurement, it is much wiser to supplement behavioural measurement with the parent-report technique. Parent-reports are widely used in research on children's vocabulary (Jørgensen, Dale, Bleses, & Fenson, 2010). Dale, Bates, Reznick, and Morisset (1989) argued that the validity and reliability of this kind of tool could be high if: (1) the parent is asked about their child's current (rather than previous) abilities or behaviour; (2) the assessed abilities have been just acquired; (3) the parent has to recognize, rather than report, the words or behaviours of their child. We assume that the CSUS follows these three criteria.

Tahiroglu et al. (2014) ran 3 studies and used the CSUS to test parents of 682 children aged 3–6 ($N_1 = 465$, $N_2 = 94$, $N_3 = 123$). The 42-item full version of the CSUS turned out to be reliable as the alphas were high (.94, .91, .90 respectively). As far as the content and constructive validity of the scale are concerned, the CSUS correlated moderately with behavioural ToM tasks, (.31 ($n_1 = 81$ children), .42 and .43 ($n_3 = 92$) respectively) and age (.47, .29, .27 respectively). In addition, correlations with the behavioural ToM tasks remained average, even when controlling for age (in study 2; $r = .35$), and for age, working and prospective memory and planning (in study 3, $r = .31$). It is also worth mentioning that 6 subscales of the CSUS (knowledge, belief, perception, intentions, desire, and emotional understanding) correlated strongly and the factor analysis revealed a one-factor structure which explained 32% of the variance. Moreover, Tahiroglu et al. (2014) pointed out in their conclusions that the relation between children's language abilities and their ToM skills measured with the CSUS should be explored.

There is no doubt that ToM abilities in preschool children are related to their language abilities (Astington & Baird, 2005). The results of a meta-analysis provided by Milligan et al. (2007) suggested that even with age controlled for, 12% of the variance in ToM abilities is explained by word comprehension – the most basic language skill. Moreover, in the Polish language it was demonstrated that word comprehension skills in 2.5-year-olds predicted ToM abilities at the ages of 3.5 and 5.5 (Białecka-Pikul, 2012). Thus, we expect that word comprehension is related to ToM abilities and if the CSUS is a valid measure of ToM, the results of the CSUS should correlate to word comprehension skills.

The main aim of our study is to adapt the CSUS to the Polish language and provide data on the psychometric properties of the CSUS among non-English-speaking children. The necessity of doing cross-cultural research on ToM development is widely accepted (e.g., Sabbagh, Xu, Carlson, Moses, & Lee, 2006;

Shahaeian et al., 2011); therefore, valid and reliable tools for measuring ToM in non-English speaking samples are also urgently needed. Thus, we took the opportunity of collecting data on the Polish version of CSUS within an ongoing longitudinal project by testing 225 parents of approximately 3.5-year-old children. This narrower age range seems plausible as, just before their fourth birthday, children are able to complete a range of different ToM tasks and therefore the CSUS can be reliably validated. Our second aim was to uncover the role of word comprehension in ToM development, measured using the CSUS.

Method

Participants

Two hundred and twenty-five caregivers (80.44% mothers) of children aged between 3 and 4 years ($M = 44.63$ months, SD = 1.47, range 38.48–47.76) participated in the study and completed the Polish version of the CSUS online. Their children attended two laboratory sessions. The final sample is presented in Table 1; detailed information regarding missing data and other characteristics of the group is available in Appendix 1 in the online Supplemental data.

Materials and procedure

This study was part of a larger three-year longitudinal research project 'The birth and development of mentalizing abilities' which took place at the Early Child Development Psychology Laboratory at the Jagiellonian University in Krakow and received clearance from the institutional ethics board. The caregivers and children were recruited on a voluntary basis via personal advertisements. Informed consent was received from all the caregivers, and the children received a small gift at the end of the lab session. During the lab session at age 3, the children were tested with 28 tasks and word comprehension test was used as the final task. During the lab session at age 3.5, ToM was tested behaviourally with 5 tasks which were presented in a fixed order in the middle of a session consisting of 22 tasks. Before or after this session, each parent received a login and registered a private account where the CSUS scale was provided online (see Appendix 1 in Supplemental data for detailed procedure).

Table 1. Characteristics of children participating in the study.

Phase of the study	Measurement	*N*			Age in months (total sample)		
		Boys	Girls	Total	Range	*M*	SD
Laboratory visit 1	Language ability	101	81	182	32.72–38.98	35.44	.52
Laboratory visit 2	Behavioural ToM	117	91	208	39.97–43.93	41.39	.44
Questionnaire	Parent-report ToM (CSUS)	126	99	225	38.48–47.76	44.63	1.47

The Polish version of the CSUS

The full version of the CSUS was used; it consists of 42 items describing children's behaviour and way of thinking (see Appendix 2 in Supplemental data for the Polish version of the scale). Parents are asked to rate their children on a 4-point Likert scale ranging from 1 (*definitely untrue of my child*) to 4 (*definitely true of my child*); a 'don't know' response was also available. The scale was translated into the Polish language and back translated by a native English speaker not acquainted with the goal of the study. All the differences between the two versions were discussed with the Polish translator and eliminated to make the Polish version as close to the original as possible.

Behavioural theory of mind tasks

Five tasks suitable for children aged 42 months were chosen to measure ToM (see Table 2 for a brief description and Appendix 3 in Supplemental data for the detailed instructions). The inter-raters' Kappa agreement coefficients for 20% of transcripts of children's answers registered on videos ranged from .79 to 1 for different tasks.

Language ability

Word comprehension was measured during the lab visit at 3 years using the Picture Vocabulary Test – Comprehension (PVTC; Haman & Fronczyk, 2012), which was constructed like the Polish version of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). It consists of 88 four-picture cards; while looking at the card, the child is asked to pick the picture that describes the word just

Table 2. Brief description of ToM tasks and percentage of children completing each task ($N = 208$).

Task	Description	% of children passing the task
Knowledge access task (Wellman & Liu, 2004)	Child judges if someone who does not see what is in the box knows the content of the box	52
Belief – emotion task (Wellman & Liu, 2004)	Child names the feeling of someone who has a mistaken belief	46
Deceptive box task (Perner, Leekam, & Wimmer, 1987)	Child judges someone's false belief about the content of the distinctive container when the child knows what is in the container	31
Modified diverse beliefs task (based on Wellman & Liu, 2004)	Child predicts where someone will search for an object based on his belief (stated explicitly), which is different from the child's belief	30
False belief task (Lazaridis, 2013)	Child judges someone's false belief that led the person to the observed behaviour	25

Note: The tasks in the Table are presented in the same order as they were administered to children.

spoken by the Experimenter. Three other pictures on the card present words that are phonetically, semantically and thematically similar. Scores could range between 0 and 88.

Analytic strategy

Missing data in the behavioural ToM tasks and word comprehension in the form of 'missing the whole task' (or the whole set of tasks in the case of ToM) were excluded from the analysis. Missing data in the CSUS only took the form of 'don't know' responses and were replaced by the mean value in the reliability and exploratory factor analysis (EFA), but was not taken into account when computing each child's final score.

Firstly, descriptive statistics were calculated. Then, the correlations between the 6 subscales of the Polish version of the CSUS were calculated and the two models (one-factor and six-factor) were analysed using confirmatory factor analyses (CFA) to describe the structure of the Polish version of the CSUS. As none of the models appeared to fit the data, EFA was conducted in the next step. Subsequently, the reliability of the CSUS was analysed using Cronbach's alpha, as in the original study (in Appendix 4 (Supplemental data), McDonald's omega reliability coefficients are also presented). The correlations between the CSUS and behavioural measures of ToM were calculated to assess its validity. Finally, correlation and regression analysis were used to explore the role of language for ToM development measured with the CSUS. All analyses were conducted using PS IMAGO PRO version 4, Mplus v. 7.4 Demo software, and JASP.

Results

Descriptive data

The mean score of the full 42-item CSUS was 3.06, with SD = .37 (range 1.85–3.88); the mean word comprehension score was 22.52, with SD = 11.72 (range 1–61). As the behavioural ToM tasks were zero-one, the percentage of children who succeeded in each task was calculated (see Table 2).

Using CFA, we revealed the one latent variable behind all five tasks measuring ToM ($\chi^2 = 3.04$, $df = 5$, $p = .69$; RMSEA = .00 [0; .07], CFI = 1); the factor score was calculated for children who completed at least two tasks ($N = 208$); the scores on missing tasks were prorated using full information maximum likelihood. The mean factor score was .02 with SD = .30 (range –.36 to .67). The ToM factor score was used in the following analyses.

The age of children did not correlate significantly with the CSUS score, word comprehension, or behavioural ToM factor score ($rs = .08$, –.02 and .11 respectively, $ps > .05$); therefore, it was not included in the subsequent analyses.

The structure of the Polish version of the CSUS

Both the one-factor and six-factor models as obtained in the original study (Tahiroglu et al., 2014) were analysed in the Polish sample using the CFA. Neither of the models had satisfactory goodness of fit indicators (for 1-factor model: $\chi^2 = 1664.17$; $df = 819$; $p < .001$; RMSEA = .067 [.063; .072]; CFI = .696; for the 6-factor model: $\chi^2 = 1601.00$; $df = 804$; $p < .001$; RMSEA = .066 [.061; .071]; CFI = .714). Therefore, the EFA was conducted in the next step.

In the Polish version of the CSUS, the mean correlation between subscales was $r = .55$, with the range of rs from .33 to .68, ps < .001, suggesting that there may not be a single factor. An EFA resulted in a two-factor solution, explaining in total 27.13% of the variance in the CSUS. Details regarding the EFA procedure and factor loadings are presented in Appendix 5 (Supplemental data). The first factor consisted of 19 items describing what the child says about the mental world and it can therefore be called Mental State Talk (MST). The second factor consisted of 13 items describing the child's understanding of mental states and can therefore be called Mental State Comprehension (MSC).

The two factors turned out to be highly correlated, with $r = .65$ ($p < .001$). The two factors differed with regard to the mean score ($t = 10.06$, $p < .001$), as children scored higher in the MST ($M = 3.25$, SD = .49) than in the MSC ($M = 2.98$, SD = .47).

The reliability of the Polish version of the CSUS

Regarding the reliability of the CSUS, Cronbach's alpha for the full scale was .90 (95% C.I. .90; .93; see Appendix 4 in Supplemental data for McDonald's omega coefficient). However, the range of item-total correlations was very wide, as the Pearson rs varied from −.06 to .66, with an average of .42 (see: Appendix 6 in Supplemental data). The reliability of the two extracted factors was also satisfactory: for the MST α = .91 (95% C.I. .90; .93; the mean item-total correlation for 19 items was $r = .55$, range .39–.68), and for the MSC α = .84 (95% C.I. .80; .87; the mean item-total correlation for 13 items was $r = .49$, range .22–.60).

Validity of the Polish version of CSUS

The behavioural ToM factor score correlated significantly, although only moderately with the CSUS total score ($r = .35$, $p < .001$), as well as with MST and MSC factors (rs = .36 and .21, both ps < .01); all these correlations remained significant when word comprehension was controlled for (rs = .33, .33 and .21 respectively, all ps < .01). Specifically, the CSUS was correlated with 4 out of 5 behavioural ToM tasks (see Appendix 7 in Supplemental data).

The role of language in ToM ability

Only the MST factor and behavioural ToM factor score correlated significantly, although weakly, with the language measure (*r*s = .19 and .20 respectively, both *p*s < .01). Correlations between word comprehension and the MSC factor, as well as the full-scale score, were not significant (*r*s = .03 and .13 respectively, both *p*s > .05). Word comprehension explained 4% of the MST factor of the CSUS ($\beta = .20$, $p < .01$); however, when behavioural ToM was introduced to the regression in the first step, word comprehension remained only marginally significant ($\beta = .12$, $p = .10$; see Appendix 8 in Supplemental data for details).

Discussion

The Polish version of the CSUS proved to be a reliable and valid measure of ToM in non-English-speaking children aged approximately 3.5 years.

Firstly and most importantly, we revealed the two-factor structure of the Polish version of CSUS, as opposed to the single-factor structure in the American sample (Tahiroglu et al., 2014). These two factors explained 27% of the variance in the CSUS, therefore distinguishing them seems to be of potential use, even though they turned out to be highly correlated. Taking into account the narrow age range of children in the study, we can suppose that the items composing the MST factor described behaviours which are easily observed in 3.5-year-olds (Bartsch & Wellman, 1989). The MSC factor's items seem more difficult for parents since assessing them requires a third-order inference, which is very cognitively demanding (Dunbar, 2014). Therefore, when assessing children's mental talk, parents can be more reliable and closer to the child's actual abilities than when assessing mental state comprehension. However, the structure of the CSUS should be further confirmed based on a more age-differentiated population.

Secondly, the Polish version of the CSUS showed high internal consistency, much like the original version (Tahiroglu et al., 2014). This was true for both the full scale and the two extracted factors, although Cronbach's alpha was smaller for the MSC factor than for MST, which can be due to the smaller number of items in MSC.

Thirdly, as far as validity of the CSUS is concerned, the parents' assessment was significantly correlated with the children's ToM performance. Even though the observed correlations were only moderate, they were in the predicted positive direction and remained significant even when language was controlled for; therefore, they do not necessarily prove the low validity of the Polish version of the CSUS. Three explanations are important here: (1) the 42-item CSUS covers a broader range of ToM abilities than lab tasks, even if we had a battery of 5 tasks; (2) some reported abilities are still developing and can only sometimes be observed in children's behaviour, in contrast to lab tasks, which are therefore more prone to chance; (3) the ecological validity of the set of tasks we used in

the lab might be disputable or related to cultural differences (see Shahaeian et al., 2011).

Fourthly, we expected that if the CSUS is a valid measure of ToM, its results should correlate to word comprehension skills. This was true only for the MST factor: the correlation was weak and the word comprehension explained only 4% of variance in ToM, assessed with the CSUS. Although this does not seem to be a strong result, it is interesting that children who understand more words may also be perceived by their parents as being more competent at talking about mental states. More generally, this raises the question about the role of children's language abilities in parent-report measures.

Concerning the limitations of our study, we have to emphasise that the CSUS was designed for children aged 3–6, and a broader age range should be adopted in the next study to validate the observed structure of the CSUS and to generalize the conclusions to preschool children. Moreover, the test–retest reliability of the Polish version of the CSUS should be addressed. Additionally, a more socioeconomically diverse sample would be of great value in future research. Further research on aspects of language other than word comprehension – especially children's mental state talk, which develops during preschool years – is also needed.

To conclude, it should be noted that the Polish version of the CSUS enabled the observation of individual differences in children's ToM abilities at the age of approximately 3.5., i.e., just about the time when they start to understand false beliefs. The Polish version of the CSUS proved to be a reliable and valid tool for studying children's abilities, as perceived by parents. The obtained results suggest that the CSUS could be a valuable tool suitable for making more thorough cross-cultural comparisons also with children speaking languages other than English. Moreover, the role of child's language in this parent-report tool was proved; therefore, it should be considered in future research with the CSUS.

Acknowledgments

We express our gratitude to all children and parents who participated in the study, as well as to all research assistants who helped with testing and data coding: Marta Szpak, Joanna Jakubowska, Martyna Jackiewicz and other team members. Special gratitude to Elżbieta Ślusarczyk, who inspired us start working on the translation of CSUS. We would also like to thank prof. Ewa Haman for sharing with us with her word comprehension test. We are also grateful to Dr Denis Tahiroglu for her valuable comments of the very first draft of this text and for Dr Takuya Yanagida for his very helpful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Polish Science Centre [grant number 2011/01/B/HS6/00453], [grant number 2015/19/B/HS6/01252] awarded to Marta Białecka-Pikul.

References

Astington, J. W., & Baird, J. A. (Eds.). (2005). *Why language matters for theory of mind*. New York:Oxford University Press.

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*(4), 946–964. doi:10.2307/1131035

Białecka-Pikul, M. (2012). *Narodziny i rozwój refleksji nad myśleniem* [*The birth and development of reflection on thinking*]. Krakow, Poland:Wydawnictwo UJ.

Dale, P. S., Bates, E., Reznick, J. S., & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language, 16*(2), 239–249. doi:10.1017/S0305000900010394

Dunbar, R. (2014). *Human evolution*. London: Penguin Books.

Dunn, L. M., & Dunn, D. M. (1997). *PPVT-III: Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.

Haman, E., & Fronczyk, K. (2012). *Obrazkowy Test Słownikowy – Rozumienie (OTSR)* [*Picture Vocabulary Test – Comprehension* (*PVT-C*)]. Gdańsk: Pracownia Testów Psychologicznych i Pedagogicznych.

Hughes, C., & Devine, R. T. (2015). Individual differences in theory of mind from preschool to adolescence: Achievements and future directions. *Child Development Perspectives, 9*, 149–153. doi:10.1111/cdep.12124

Jørgensen, R. N., Dale, P. S., Bleses, D., & Fenson, L. (2010). CLEX: A cross-linguistic lexical norms database. *Journal of Child Language, 37*(2), 419–428. doi:10.1017/S0305000909009544

Lazaridis, M. (2013). The emergence of a temporally extended self and factors that contribute to its development: From theoretical and empirical perspectives. *Monographs of the Society for Research in Child Development, 78*(2), 1–120. doi:10.1111/mono.v78.2/issuetoc

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622–646. doi:10.1111/j.1467-8624.2007.01018.x

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125–137. doi:10.1111/j.2044-835X.1987.tb01048.x

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind a comparison of Chinese and US preschoolers. *Psychological Science, 17*(1), 74–81. doi:10.1111/j.1467-9280.2005.01667.x

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology, 47*(5), 1239–1247. doi:10.1037/a0023899

Tahiroglu, D., Moses, L. J., Carlson, S. M., Mahy, C. E., Olofson, E. L., & Sabbagh, M. A. (2014). The children's social understanding scale: Construction and validation of a parent-report measure for assessing individual differences in children's theories of mind. *Developmental Psychology, 50*(11), 2485–2497. doi:10.1037/a0037914

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541. doi:10.1111/j.1467-8624.2004.00691.x