The title

Karolina Muszyńska[1] & XXX[2]

[1] University of Warsaw, Faculty of Psychology

[2] Stanford University

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Karolina Muszyńska: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; XXX: Writing - Review & Editing, Supervision.

Correspondence concerning this article should be addressed to Karolina Muszyńska, University of Warsaw, Faculty of Psychology, ul. Stefana Banacha 2D, 02-097 Warsaw, Poland. E-mail: karolina.muszynska@psych.uw.edu.pl

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline. Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines. One sentence clearly stating the **general problem** being addressed by this particular study. One sentence summarizing the main result (with the words "**here we show**" or their equivalent). Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge. One or two sentences to put the results into a more **general context**. Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: keywords

Word count: X

The title

Methods

(to paste from google doc)

Participants

(to paste from google doc) Mention the bilinguals and multilinguals

Material

(to paste from google doc)

Procedure

(to paste from google doc)

Data analysis

Results

Psychometric properties of the two CAT-CDIs

Our first aim was to examine whether CAT-CDIs in American English and Polish demonstrate comparable psychometric properties. To that end, we revisit the psychometric properties reported for the American English CAT-CDI (word production) in Kachergis et al. (2022) and compare those to the data from Polish CAT-CDI (Words and Sentences).

We found similarly strong correlations in the two languages between the abilities estimated from CDI-CAT and full CDI scores (American English and Polish: $r = .86$), the abilities estimated from the CDI-CAT and abilities estimated from full CDI (American

Table 1

American English: Correlations between ability estimated by CAT-CDI and ability estimated from full CDI by children's age

	[15,18)	[18,21)	[21,24)	[24,27)	[27,30)	[30,33)	[33,36]
r ability CAT vs full CDI	0.95	0.85	0.82	0.83	0.59	0.84	0.86
N	26	22	26	30	28	24	48

English and Polish: $r = .92$), and the abilities estimated from the full CDI and the full CDI scores (American English: $r = .95$, Polish: $r = 0.94$). The abilities estimated from the CDI-CAT and the full CDI scores were also strongly correlated within individual age groups (see Table 2).

The Polish validation study included 28 data from bi- and multilingual families. Though it is a small group, we decided to explore their correlation coefficients (non-parametric Spearman's rho) and found these were similar to those found for Polish monolingual children (see Table 3 in Supplementary Materials).

We also compared the mean squared error between the abilities as estimated by CAT-CDI and from the full CDI. The mean squared error in English was 0.55 ($Mdn = 0.17$, $SD = 1$), and in Polish it was 0.19 ($Mdn = 0.08$, $SD = 0.45$). We also looked at the children for whom the estimates from the CAT-CDI and full CDI diverged extremely, i.e. their difference between the errors was 1.5 SD from the mean. There were 15 such cases (7.35%) in the English dataset and 4 cases (1.96%) in the Polish dataset. All participants (in both datasets) showed a higher CDI-CAT ability as compared to their ability estimated from the full CDI. All of the Polish participants showing extreme discrepancy completed the full CDI in an unusually short time–their completion times were among the shortest 5% in the sample, meaning 95% of participants took longer. This might suggest their ability estimates from the full CDI may have been underestimated, possibly due to rushed

Table 2

Polish: Correlations between ability estimated by CAT-CDI and ability estimated from full CDI by children's age

	[18,21)	[21,24)	[24,27)	[27,30)	[30,33)	[33,36]
r ability CAT vs full CDI	0.8	0.94	0.91	0.89	0.95	NA
N	29	22	16	23	22	1

(and/or less attentive responses). In support of this view, their CAT-CDI scores had acceptable measurement errors (<= 0.15 for English, <= 0.1 for Polish), indicating that the CAT-CDI was able to estimate their abilities reliably, in contrast to the full CDI. However, the pattern was not replicated in the English dataset, where only 2 participants who showed extreme discrepancy also showed very short administrations of the full CDI.

NOTE TO US: Below I'm printing the participants with extreme discrepancies, you can see they vary in age, duration of CAT/full, SE of theta, etc. I myself don't see a pattern here:

We also calculated the mean squared error without the cases of extreme discrepancy, which yielded a MSE of XX in English and MSE of XX in Polish.

Discussion

1. Correlations strong in both languages (overall and in age-bins; in PL: multi and mono correlations similar and strong).
2. MSE after removing cases with extreme discrepancies.
3. Extreme discrepancies - in both lgs, we see XXX

Table 3

Supplementary Material: Table S1 - Spearman's correlations for monolingual and multilingual children in the Polish dataset

lang_group	r	n	correlation
monolingual	0.92	85	Ability from CDI-CAT ~ full CDI score
multilingual	0.90	28	Ability from CDI-CAT ~ full CDI score
monolingual	0.92	85	Ability from CDI-CAT ~ ability from full CDI
multilingual	0.90	28	Ability from CDI-CAT ~ ability from full CDI
monolingual	1.00	85	Ability from full CDI ~ full CDI score
multilingual	1.00	28	Ability from full CDI ~ full CDI score

References

Auguie, B. (2017). *gridExtra: Miscellaneous functions for "grid" graphics*. Retrieved from https://CRAN.R-project.org/package=gridExtra

Aust, F., & Barth, M. (2024). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://doi.org/10.32614/CRAN.package.papaja

Barth, M. (2023). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Chalmers, R. P. (2016). Generating adaptive and non-adaptive test interfaces for multidimensional item response theory applications. *Journal of Statistical Software*, *71*(5), 1–39. https://doi.org/10.18637/jss.v071.i05

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., … Borges, B. (2024). *Shiny: Web application framework for r*. Retrieved from

Table 4

production	sex_full	age_full	order	fullTheta	fullTheta_SE	catTheta	catTheta_SE	sq_
97.00	Female	27.00	full_first	-0.14	0.04	1.20	0.17	1.81
8.00	Male	17.00	cat_first	-1.58	0.16	-0.23	0.16	1.82
158.00	Male	35.00	full_first	0.14	0.04	1.62	0.16	2.17
0.00	Male	34.00	full_first	-2.90	0.43	-1.48	0.38	2.01
132.00	Female	21.00	cat_first	0.02	0.04	1.75	0.17	2.99
165.00	Male	20.00	full_first	0.18	0.04	1.48	0.17	1.71
47.00	Female	28.00	full_first	-0.57	0.06	1.86	0.17	5.90
14.00	Male	20.00	full_first	-1.27	0.12	0.01	0.16	1.64
124.00	Female	28.00	cat_first	0.00	0.04	1.30	0.17	1.68
210.00	Female	26.00	cat_first	0.33	0.03	1.85	0.17	2.31
5.00	Female	26.00	cat_first	-1.79	0.19	-0.42	0.19	1.87
177.00	Male	28.00	full_first	0.22	0.03	1.62	0.18	1.98
470.00	Male	36.00	full_first	1.14	0.03	2.82	0.35	2.83
253.00	Male	35.00	cat_first	0.48	0.03	1.83	0.16	1.83
287.00	Male	23.00	full_first	0.58	0.03	1.91	0.16	1.78

https://CRAN.R-project.org/package=shiny

Garnier, Simon, Ross, Noam, Rudis, Robert, ... Cédric. (2023). *viridis(Lite) - colorblind-friendly color maps for r*. https://doi.org/10.5281/zenodo.4678327

Garnier, Simon, Ross, Noam, Rudis, Robert, ... Cédric. (2024). *viridis(Lite) - colorblind-friendly color maps for r*. https://doi.org/10.5281/zenodo.4679423

Grolemund, G., & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, *40*(3), 1–25. Retrieved from https://www.jstatsoft.org/v40/i03/

Kassambara, A. (2023). *Ggpubr: 'ggplot2' based publication ready plots*. Retrieved from https://CRAN.R-project.org/package=ggpubr

Krajewski, G. (2025). *Multilada: MultiLADA's little helpers*. Retrieved from https://github.com/gkrajewski/Multilada

Müller, K. (2020). *Here: A simpler way to find your files*. Retrieved from https://CRAN.R-project.org/package=here

Müller, K., & Wickham, H. (2025). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

R Core Team. (2025). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org

Slowikowski, K. (2024). *Ggrepel: Automatically position non-overlapping text labels with 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggrepel

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2023a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2023b). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2025). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

Wickham, H., Hester, J., & Bryan, J. (2024). *Readr: Read rectangular text data*. Retrieved from https://CRAN.R-project.org/package=readr

Wickham, H., Vaughan, D., & Girlich, M. (2024). *Tidyr: Tidy messy data*. Retrieved from https://CRAN.R-project.org/package=tidyr

R (Version 4.4.3; R Core Team, 2025) and the R-packages *dplyr* (Version 1.1.4; Wickham, François, Henry, Müller, & Vaughan, 2023), *forcats* (Version 1.0.0; Wickham, 2023a), *ggplot2* (Version 3.5.2; Wickham, 2016), *ggpubr* (Version 0.6.0; Kassambara, 2023), *ggrepel* (Version 0.9.6; Slowikowski, 2024), *gridExtra* (Version 2.3; Auguie, 2017), *here* (Version 1.0.1; Müller, 2020), *lattice* (Version 0.22.7; Sarkar, 2008), *lubridate* (Version 1.9.4; Grolemund & Wickham, 2011), *mirt* (Version 1.44.0; Chalmers, 2012, 2016), *mirtCAT* (Version 1.14; Chalmers, 2016), *Multilada* (Version 0.7.0; Krajewski, 2025), *papaja* (Version 0.1.3; Aust & Barth, 2024), *purrr* (Version 1.0.4; Wickham & Henry, 2025), *readr* (Version 2.1.5; Wickham, Hester, & Bryan, 2024), *shiny* (Version 1.10.0; Chang et al., 2024), *stringr* (Version 1.5.1; Wickham, 2023b), *tibble* (Version 3.3.0; Müller & Wickham, 2025), *tidyr* (Version 1.3.1; Wickham, Vaughan, & Girlich, 2024), *tidyverse* (Version 2.0.0; Wickham et al., 2019), *tinylabels* (Version 0.2.5; Barth, 2023), *viridis* (Garnier et al., 2023; Version 0.6.5; Garnier et al., 2024) and *viridisLite* (Version 0.4.2; Garnier et al., 2023)