



GUIDANCE FOR USING ONLY SPECIFIC CREDI DOMAINS

October 17 , 2018

Jorge Cuartas
Harvard Graduate School of Education

Sensitivity analyses for CREDI scores with missing domains

Q: I want to use the CREDI Long Form, but I am only interested in one specific domain (for example, social-emotional development). Can I only use items that correspond to that domain?

A: Yes, but with important caveats. Under normal circumstances, the scoring procedure for the CREDI Long Form takes into account information from all domains when producing its individual scores for the motor, language, cognitive, and social-emotional domains. As such, there are several limitations to producing scores using anything less than the full Long Form tool. First, having missing items would significantly reduce the sample that can be scored using the CREDI package. Second, even though the scores produced when items from some domains are missing are highly correlated and show strong agreement with the scores using the full Long Form, our analyses suggest that they are not exactly the same. Particularly, as more domains are missing (i.e., items from other domains that do not overlap with the domain of interest), there is a greater discrepancy between domain scores, with differences that can be up to 0.3 SD relative to the scores that would have been produced with the full Long Form. As such, to maximize precision, we highly encourage the inclusion of all items from all domains when using the CREDI. Nevertheless, if using the full CREDI Long Form is not of interest or practical in your study, you may focus on items from a more limited set of domains, recognizing that their scores are likely to contain some error. (Please see *CREDI Long Form Item Domain Mapping* document for information on which Long Form items correspond to which domain.) If you do use a limited set of items, please note that you should never trust the scores produced by the Long Form's scoring system for the domains that were not included in your original data collection.

1. Introduction

In this document we present sensitivity analyses for the scoring of the CREDI socioemotional domain when other domains are missing. Specifically, we estimate the CREDI long form scores using all the domains and excluding all the items that did not overlap with the socioemotional domain from 1) the motor domain, 2) the cognitive domain, 3) the language domain, 4) the motor and cognitive domains, and 5) the motor, cognitive, and language domains.

To analyze the extent to which socioemotional scores differ according to missingness of one or more domains, we employed different analytic strategies. First of all, we analyze the differences between socioemotional raw scores and standardized scores using mean difference tests. Second, to account for the fact that raw scores increase with age, reflecting developmental progression (McCoy, Fink, Waldman, 2018), we performed regression models regressing the score predicted with the missing domain(s) and the score produced including all domains against different functional forms of age (linear, quadratic, and cubic). Residuals from these models reflect the component in scores not accounted for by age, which were subsequently used to perform mean difference tests. Third, we performed regression models to analyze the correlation between scores, using different specifications including different functional forms of age. Fourth, we performed non-parametric local weighted scatterplot smoothing (LOWESS) to account for possible non-linearities in the relation between scores, and to analyze the relation between scores without assuming any functional form to fit the model. Finally, we estimated Lin's concordance correlation coefficient (Lin, 1989) and Bland-Altman limits of agreement (Bland & Altman, 1986) to analyze the extent to which scores produced including all against some domains agree. Findings suggest that socioemotional raw scores are relatively robust to the exclusion of other domains. Yet, the scores produced when there are missing domains are not statistically identical

to those including all domains, and as more domains are missing the disagreement between scores grows, so the lack of domains may compromise the precision of the estimated scores, especially as more domains are missing. Moreover, the missingness of domains (and items) may reduce the number of cases that the CREDI package can score, compromising the final sample that is scored. The loss in precision and sample may be especially problematic when missing more than one dimension. As such, we highly encourage the inclusion of all domains when using the CREDI, but if it is not feasible and you are interested exclusively in one domain, we encourage the use of as many domains as possible to maximize the scoring precision and the sample size that can be scored.

2. Mean difference test between scores

To begin, Table 1 presents the mean score for the original measurement and the measurement missing domains for the sample that was scored in the latter, as well as the raw difference between scores. As shown, the difference between scores is statistically significant in all cases but is relatively small in magnitude: it lies between -0.004 units to -0.100 on average, which corresponds to averages differences of -0.001 SD to -0.040 SD of the original score. In this table it is also important to note that there are considerable reductions in the sample that can be scored as more items are missing. Results are robust when using residuals from regression models after controlling for age.

Table 1. Mean difference tests

Missing domain	Mean		Difference		N
	Original scores	Missing domains	Raw scores	Sig.	
1. Motor	50.547	50.543	-0.004	**	12,008
2. Cognitive	50.332	50.319	-0.012	**	14,004
3. Language	50.178	50.173	-0.005	**	12,993
4. Motor + Cognitive	50.591	50.555	-0.036	**	11,987
5. Motor + Cognitive + Language	50.557	50.457	-0.100	**	10,116

** $p < 0.01$, * $p < 0.05$

3. Correlation between scores

Table 2 presents the correlation between scores produced by the original measurement and by a measurement with missing domains. As can be seen, the correlation is above 0.99 on all cases, but is smaller as more items are missing. On the other hand, results from regression models further show that there is a relation that is almost 1-to-1 between scores, even when controlling for different functional forms of age.

Table 2. Correlation and regression models

Missing domain	Correlation	Estimated coefficient				N
		(1)	(2)	(3)	(4)	
1. Motor	0.999	1.000 (0.000)	0.990 (0.000)	0.994 (0.001)	0.994 (0.007)	12,008
2. Cognitive	0.998	1.009 (0.000)	1.014 (0.001)	1.015 (0.001)	1.014 (0.001)	14,004
3. Language	0.999	0.998 (0.000)	0.998 (0.003)	0.995 (0.000)	0.995 (0.000)	12,993
4. Motor + Cognitive	0.996	1.018 (0.008)	1.006 (0.002)	1.001 (0.002)	1.002 (0.003)	11,987
5. Motor + Cognitive + Language	0.992	1.026 (0.001)	0.988 (0.002)	0.976 (0.004)	0.980 (0.004)	10,116

Standard errors in parentheses. (1) bivariate relation, (2) including age (3) quadratic term of age (4) cubic term of age. $R^2 > 0.99$ in all models.

4. LOWESS curves

Figures 1-5 present LOWESS curves between the original score and scores produced when different domains are missing. As can be seen, even relaxing the linearity assumptions of regression models, there is a strong relation between scores. Yet, as more domains are missing, the dispersion around the fitted curve increases.

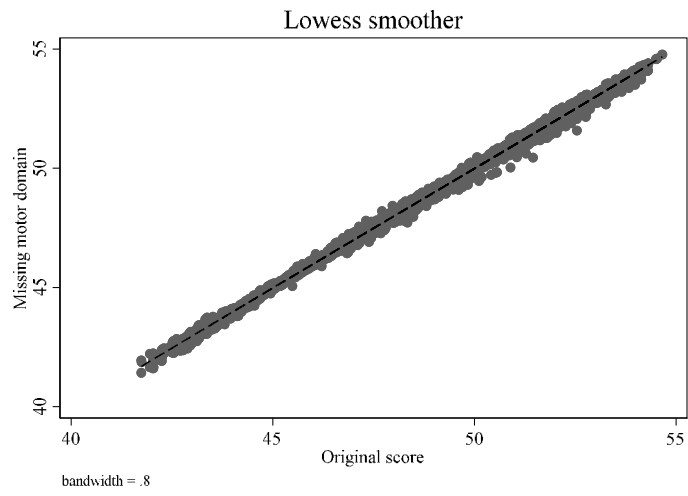


Fig 1. LOWESS curve for score without motor domain

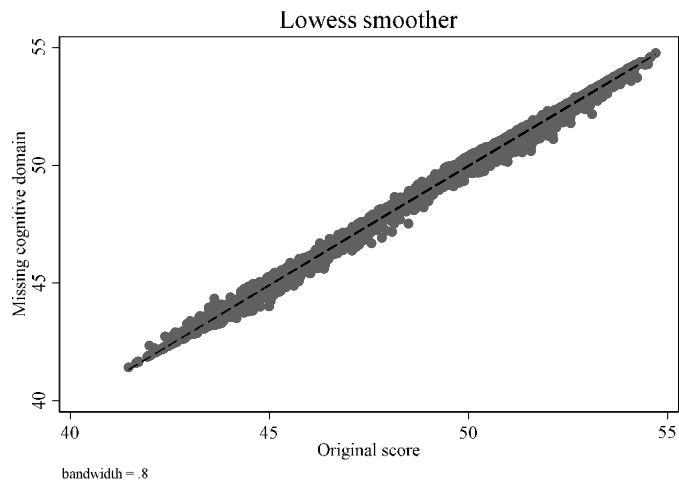


Fig 2. LOWESS curve for score without cognitive domain

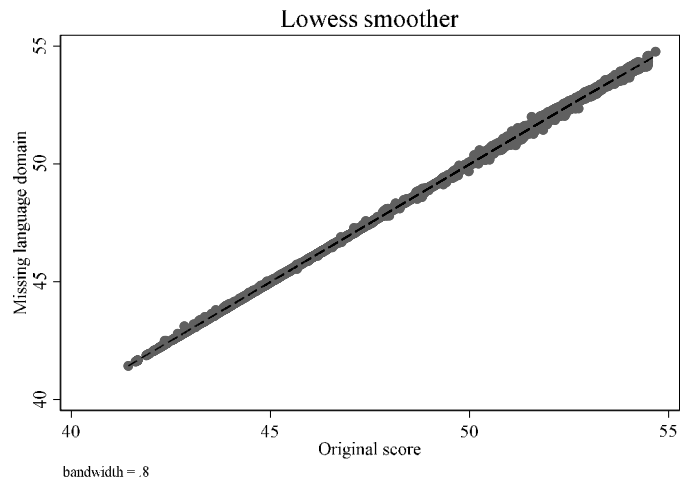


Fig 3. LOWESS curve for score without language domain

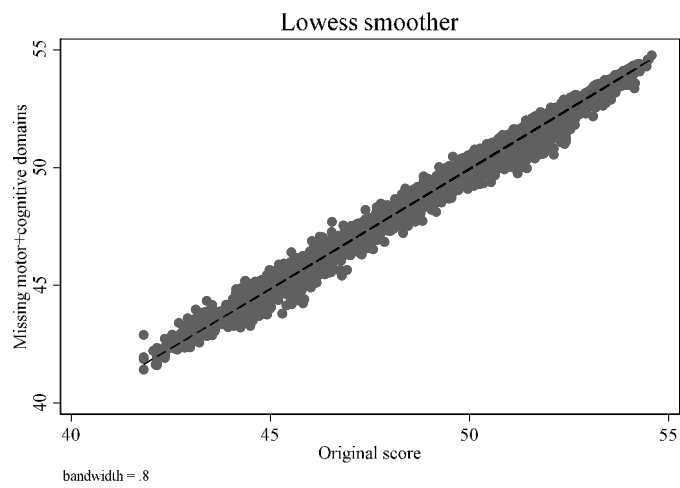


Fig 4. LOWESS curve for score without motor and cognitive domains

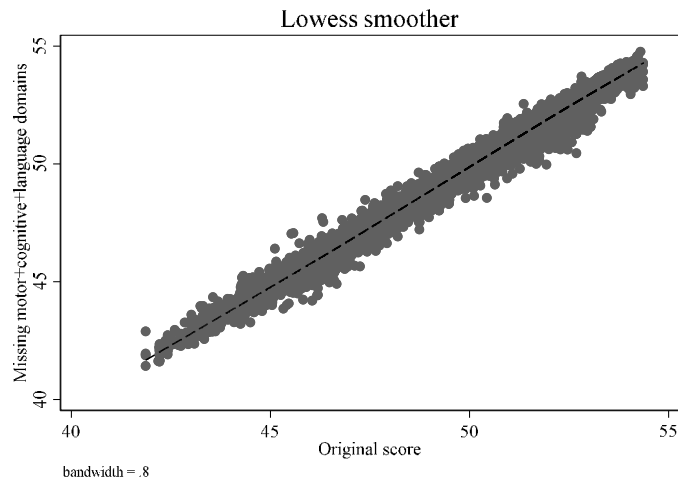


Fig 5. LOWESS curve for score without motor, cognitive, and language domains

5. Lin's concordance correlation coefficient and Bland-Altman limits of agreement (LOA)

Having a high correlation between measures does not guarantee that both measures are equal, mainly because it does not explicitly show whether the measurements fall on a 45° line (i.e., showing agreement). The concordance correlation coefficient (Lin, 1989) evaluates the agreement between two measures, examining the variation from the 45° line through the origin. Table 3 presents the concordance correlation coefficient between the scores with full information and with missing domains. As can be seen, all coefficients are above 0.99, and when the missing domain is language it is 1. According to McBride (2005) and Altman (1991), a concordance correlation can be considered as excellent if it is above 0.99, and as substantial if it is between 0.95 and 0.99.

Table 3. Strength of agreement between measures

Missing domain	Lin's concordance	Raw difference		
		Average	95% LOA	
1. Motor	0.999	-0.004	-0.192	0.184
2. Cognitive	0.998	-0.012	-0.286	0.262
3. Language	1.000	-0.005	-0.092	0.082
4. Motor + Cognitive	0.995	-0.036	-0.510	0.438
5. Motor + Cognitive + Language	0.990	-0.100	-0.747	0.547

As an additional method to measure agreement when scoring with full information and with missing domains, we estimate Bland and Altman (1986) limits of agreement (LOA). LOA show how far apart measurements can be; specifically, it estimates an agreement interval, in which 95% of differences between measurements are expected to fall, if these differences follow a normal distribution. Table 3 and Figs 6-10 presents the LOA. Results show that the LOA are wider as more domains are missing. Particularly, 95% of the differences between measurements are expected to fall between -0.092 and 0.082 when the language domain is missing (which corresponds to -0.04 SD to 0.03 SD of the original score), up to -0.747 and 0.547 when the motor, cognitive, and language domains are missing (-0.30 SD to 0.22 SD of the original score).

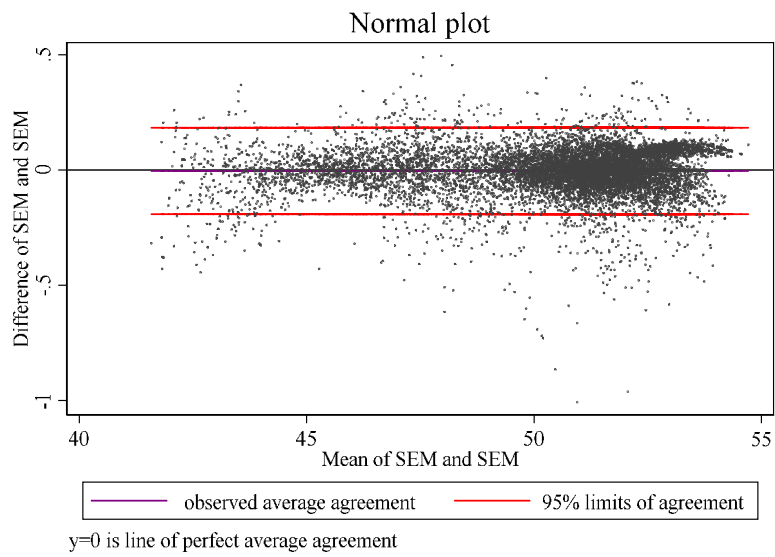


Fig 6. Limits of agreement – without motor domain

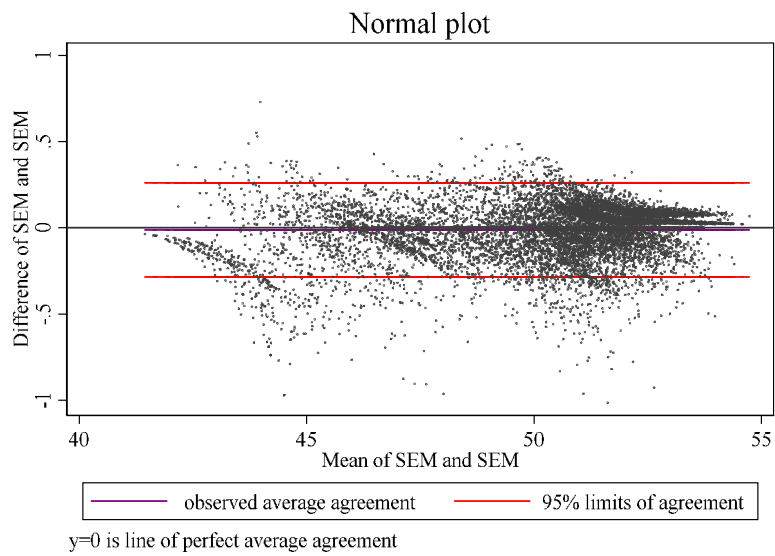


Fig 7. Limits of agreement – without cognitive domain

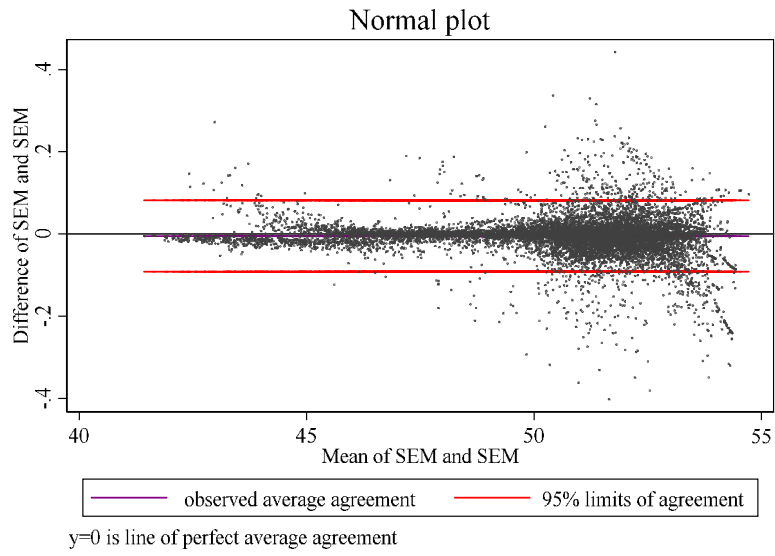


Fig 8. Limits of agreement – without language domain

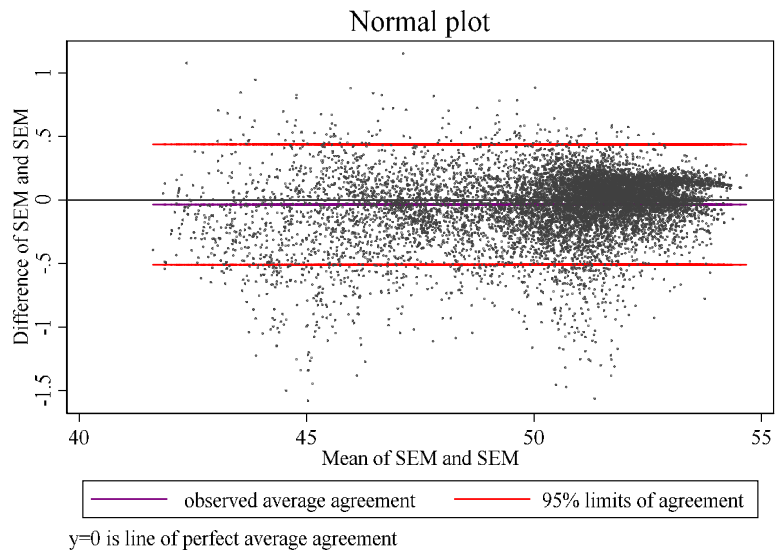


Fig 9. Limits of agreement – without motor and cognitive domains

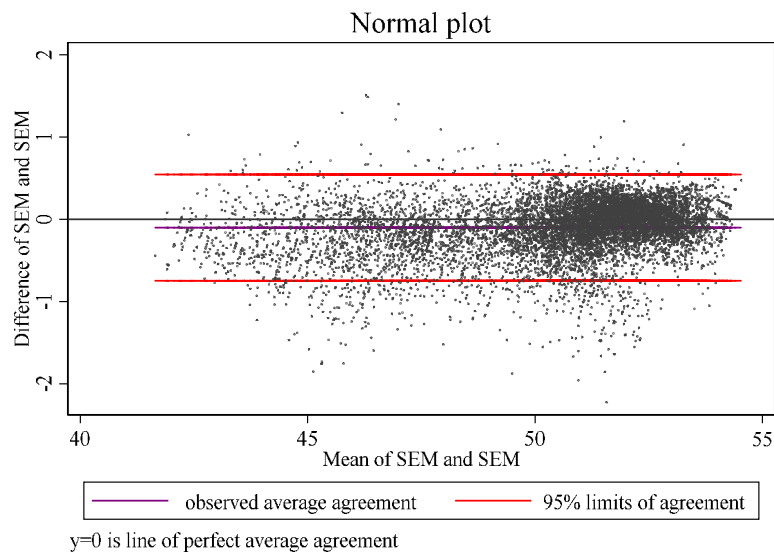


Fig 10. Limits of agreement – without motor, cognitive and language domains

6. Conclusion

In conclusion, it is possible to use the CREDI to estimate socioemotional scores even without having all the other domains. Nonetheless, it is important to note that having missing domains may compromise the precision of the estimates, and as more domains are missing the measure will be less accurate. Moreover, having missing domains may reduce the sample that is effectively scored. Given this, we highly encourage the use of all CREDI domains, but if it is not feasible, or your interest is to obtain scores for the socioemotional domain only, we encourage the use of as many domains as possible to maximize measurement precision and the sample size that is scored. Although the CREDI Long Form’s scoring system will produce scores for all domains, regardless of which items were collected, it is important to recognize that you should use the scores produced for domains that were not part of your original data collection.

7. References

- Altman, D. (1991). *Practical statistics for medical research*. London: Chapman and Hall/CRC.
- Bland, M. J., & Altman, D. (1986). Statistical Methods For Assessing Agreement Between Two Methods Of Clinical Measurement. *The Lancet*, 327(8476), 307-310.
doi:[https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Lin, L. I. K. (1989). A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics*, 45(1), 255-268. doi:10.2307/2532051
- McBride, G. B. (2005). *A proposal for strenght-of-oagreement criteria for Lin's concordance correlation coefficient*. Retrieved from Hamilton, New Zealand
- McCoy, D., Fink, G. & Waldman, M. (2018). Caregiver-reported early development instruments (CREDI) data management & scoring manual. Harvard University