High-Tech Augmentative and Alternative Communication for Individuals with IDD and Complex Communication Needs: A Meta-Analysis (**Design Quality Review**)

Jennifer B. Ganz[1], Ph.D., jeniganz@tamu.edu; Kristi Morin[2], M.Ed., klmorin@tamu.edu; Margaret J. Foster[3], MPH, margaretfoster@tamu.edu; Kimberly J. Vannest[4], Ph.D. kvannest@tamu.edu; Derya Genç Tosun[5], Ph.D., drygnc@gmail.com; Emily V. Gregori[6], M.Ed., egregori@purdue.edu; Stephanie Gerow[7], Ph.D., stephanie_gerow@baylor.edu

[1, 2, 4] Texas A&M University

4225 Texas A&M University, College Station, TX 77843-4225

[3] Texas A&M University

4462 TAMU, Medical Sciences Library, College Station, TX 77843

[5] Anadolu Universitesi

Engelliler Arastirma Enstitusu, Eskişehir 26470, Turkey

[6] Purdue University

Beering Hall, West Lafayette, IN 47907-2098

[7] Baylor University

Marrs McLean Science 316, One Bear Place #97301, Waco, TX 76798

Corresponding Author: Jennifer B. Ganz, Ph.D.

Design quality review. Each article that met the inclusion criteria for full-text review was further evaluated for inclusion related to design quality using the Basic Design Standards for single-case research proposed by What Works Clearinghouse (WWC; Kratochwill et al., 2010) and adapted by Maggin and colleagues (2013). There is a total of four standards, two sub-standards, and an overall evaluation (described further below). Any article that scored a "0" on the overall evaluation, meaning that it did not meet standards, was excluded from further evaluation. A total of 59 articles were coded for design quality; 8 of these articles did not meet basic design quality standards and an additional 20 articles were excluded because the dependent variable was not AAC, resulting in 31 articles coded for potential moderators.

WWC Design standard 1: Experimental Design. The first standard (Kratochwill et al., 2010) evaluates whether the independent variable is systematically manipulated, meaning the research design included manipulation of the independent variable such that a functional relation could be detected if there was one. If the design was such that the independent variable was systematically manipulated, the article was scored as a "1" and if the design was such that the independent variable was not systematically manipulated, the article was scored as a "0".

WWC Design standard 2a: Systematic IOA. The second standard (Kratochwill et al., 2010) is broken into three sub-standards and evaluates the extent to which the article reported inter-observer agreement (IOA) results. Design standard 2a evaluates whether more than one assessor systematically measured each outcome variable over time. In other words, it assessed whether the article reported IOA results. If the article did report IOA results, it was scored as a "1". If the article did not report IOA results, it was scored as a "0".

WWC Design standard 2b: IOA by condition. Design standard 2b (Kratochwill et al., 2010) further evaluated the extent to which the article reported IOA results by evaluating

whether IOA was collected in each condition and on at least 20% of the data points in each condition (i.e., baseline, intervention, generalization, and/or maintenance conditions). If the article reported that IOA was collected on a minimum of 20% of the data points within each condition, the article was scored as a "2". If the article reported that IOA was collected on a minimum of 20% of the data points, but the article did not specify that it was collected for a minimum of 20% of the data points for every condition, then it was scored as a "1". Lastly, if the article did not report that IOA was collected or if it reported that it was collected for less than 20% in each condition, then the article was scored as a "0". This standard was modified from the original WWC document. According to WWC, if an article does not specify that IOA was collected for a minimum of 20% of the data points within each condition, then the article does not meet standards (e.g., would be scored as a "0"). However, we found that a large number of articles collected IOA for a minimum of 20% of data points but failed to disaggregate the results by phase or condition. In order to avoid eliminating these studies from our analysis based solely on this criteria, we added the intermediary scoring criteria.

WWC Design standard 2c: IOA Results. Design standard 2c (Kratochwill et al., 2010) also evaluated the extent to which the article reported IOA by evaluating whether the reported IOA results met minimum thresholds for all phases. If the article reported IOA results that were .80 or higher for percentage agreement indices or .60 or higher for kappa measures, and the article reported IOA results for each condition (rather than an aggregated score), then the article was scored as a "2". If the article reported IOA results that were .80 or higher for percentage agreement indices or .60 or higher for kappa measures but did not disaggregate the results by condition, then the article was scored as a "1". Lastly, if the article did not report IOA results or if the results were not at least .80 or higher for percentage agreement indices or .60 or higher for

kappa measures, then the article was scored as a "0". We modified this standard from the original WWC Standard by adding the intermediary scoring criteria for the same reason as is listed in Design Standard 2b.

   WWC Design standard 3: Experimental Control. The third standard (Kratochwill et al., 2010) evaluated whether the study could demonstrate experimental control by determining if the study included at least three attempts to demonstrate an intervention effect at three different points in time or with three different phase repetitions (e.g., "ABAB designs and their extensions, multiple-baseline designs with at least three baseline conditions, changing criterion designs with at least three different criteria, and more complex variants of these designs...Alternating and simultaneous treatment design requires five opportunities to demonstrate a treatment effect and three data points" p. 28, Kratochwill et al, 2013). When determining whether or not the study demonstrated an intervention effect, only adjacent pairwise comparisons were considered (e.g., AB). If the experiment included at least three attempts to demonstrate an intervention effect at three different points in time (or at two conditions for alternating treatment designs), then the article was scored as a "1". If the experiment did not include at least three attempts to demonstrate an intervention effect at three different points in time (or didn't include at least two conditions for alternating treatment designs), then the article was scored as a "0".

   WWC Design standard 4: Phase Length. The fourth standard (Kratochwill et al., 2010) evaluated whether the article attempted to demonstrate an effect by determining if the article reported a minimum of three data points per baseline and intervention phases. For reversal or withdrawal designs, if the graph included a minimum of four phases (two baseline and two intervention phases) with at least five data points per phase, the article was scored as a "2". If the

graph included a minimum of four phases (two baseline and two intervention phases) with at least three data points per phase, the article was scored as a "1". If any phases included fewer than three data points, the article was scored as a "0". For multiple-baseline and multiple-probe designs, if the graph included a minimum of six phases (one baseline and one intervention phase for each of the three levels) with at least five data points per phase, the article was scored as a "2". If the graph included a minimum of six phases (one baseline and one intervention phase for each of the three levels) with at least three data points per phase, the article was scored as a "1". If any of the phases included fewer than three data points, the article was scored as a "0". For alternating treatment designs, if the graph included at least five repetitions of the alternating sequence (i.e., at least five data points per treatment), then the article was scored as a "2". If the graph included at least four repetitions of the alternating sequence (i.e., at least four data points per treatment), then the article was scored as a "1". If the graph included less than four repetitions of the alternating sequence (i.e., less than four data points per treatment), then the article was scored as a "0".

   Overall evaluation. After each article was evaluated according to each of the Design Standards described above, an overall score was assigned. If the article met all standards without reservations (i.e., scored a "1" on Standards 1, 2a, and 3 and scored a "2" on Standards 2b, 2c, and 4), then it was scored as a "2", meaning that the article met the standards without reservations. If the article met some standards with reservations (i.e., scored a "1" on Standards 2b, 2c, or 4) but did not score a "0" on any standard, then it was scored as a "1", meaning that the article met the standards with reservations. If the article scored a "0" on any standard, then it was scored as a "0", meaning that it did not meet the standards.

References

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., &

Shadish, W. R. (2010). *Single-case designs technical documentation.* Retrieved from

http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Maggin, D. M., Briesch, A. M., & Chafouleas, S. M. (2013). An application of the What Works

Clearinghouse standards for evaluating single-subject research: Synthesis of the self-

management literature base. *Remedial and Special Education, 34,* 44-58.

doi:10.1177/0741932511435176