



Best Practice Considerations For Assessment and Evaluation During Distance Learning

Ryan Farmer, PhD; Jackie Johns, EdS, NCSP; Kathryn Moore, PhD;



Remote Administration: What Do We Know?



Based on Farmer, McGill, Dombrowski, et al., (in press)

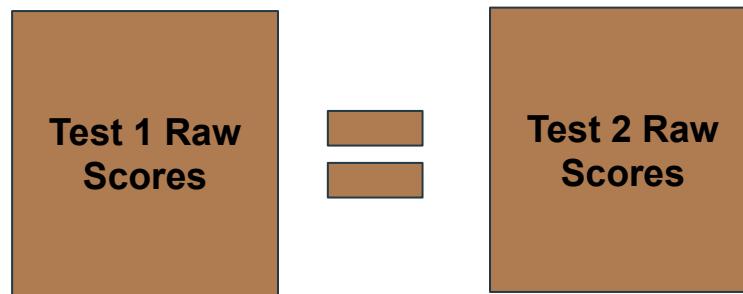
What evidence is necessary?

Like with any new test, publishers are responsible to produce reliability and validity evidence. The International Test Commission (2016) defines new formats of available tests as new versions, warranting documentation of:

- Construct Validity
- Validity from Response Processes
- Inferential validity
- Reliability
- etc

What has been provided?

Focus of evidence provided by publishers has been on “raw score equivalence.” This is all about the normative samples.



... the plural in ‘scores’ is important.

Three Relevant Studies

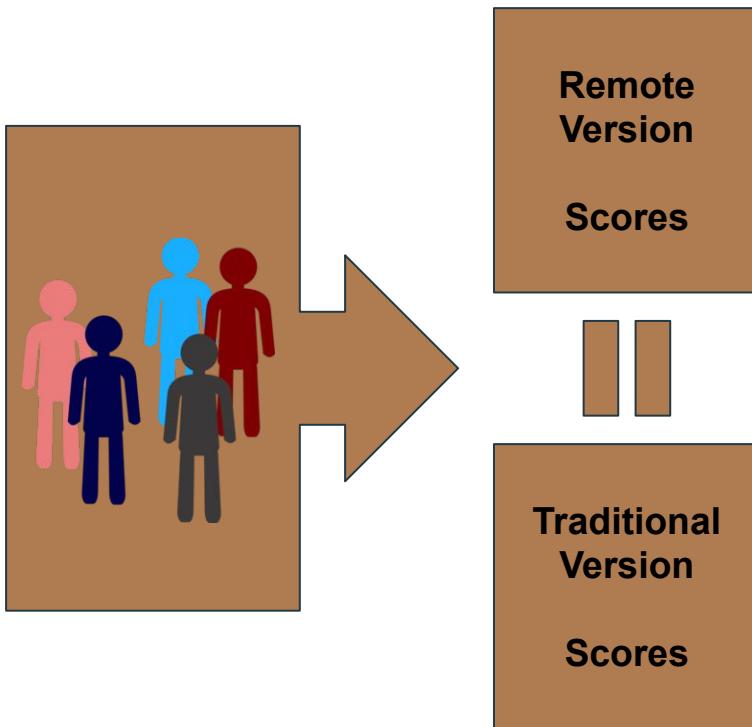
- WISC-5 -- Daniel and Wahlstrom, 2019
 - WJ IV COG/ACH -- Wright, 2018a
 - RIAS-2 -- Wright 2018b
- All studies...
- Carefully controlled the environment
 - Tested students in a clinic or school setting in all conditions (no home settings)
 - Sometimes used trained facilitators
 - Controlled technology hardware and internet
 - Used equivalent-groups design

Three Relevant Studies

- WISC-5 -- Daniel and Wahlstrom, 2019
 - WJ IV COG/ACH -- Wright, 2018a
 - RIAS-2 -- Wright 2018b
- All studies...
- Carefully controlled the environment
 - Tested students in a clinic or school setting in all conditions (no home settings)
 - Sometimes used trained facilitators
 - Controlled technology hardware and internet
 - **Used equivalent-groups designs**

This is very important!

Equivalent-Groups Design



Effectively, EG designs ask the question:

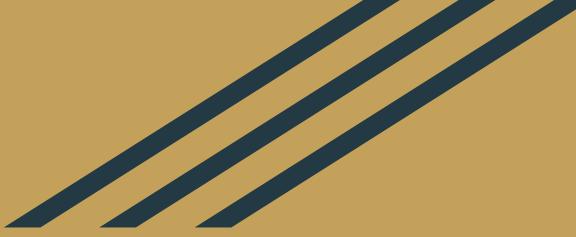
Is the mean score of group 1 statistically different from the mean score of group 2?

Pretty much across all three studies, the answer was a loud “yes!” (with a few caveats).

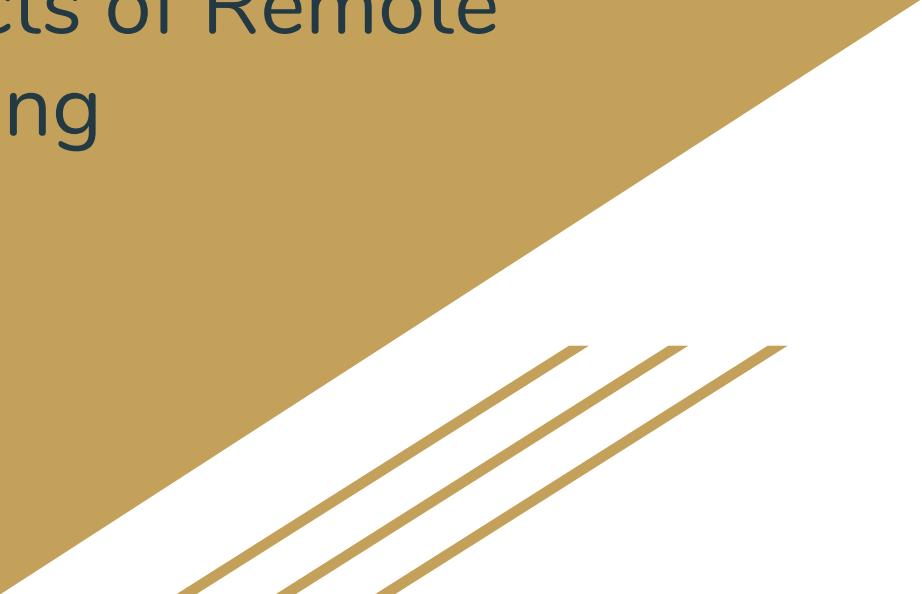
As such, the mean IQ of the Remote group was ~100. The mean IQ of the traditional group was also ~100.

Are we asking the wrong question?

Do we want to know if the group averages are statistically the same, or do we want to know if a child would get an approximately similar score regardless of format?



Practical Aspects of Remote Testing



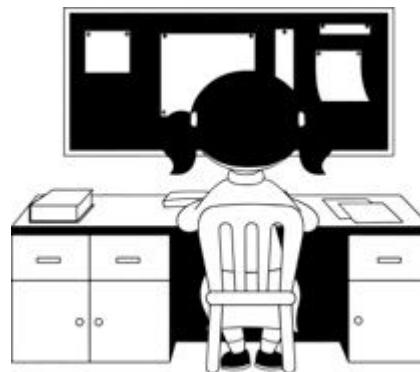
Unique Sources of Error in Remote Admin



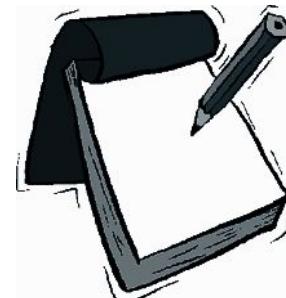
Examinee



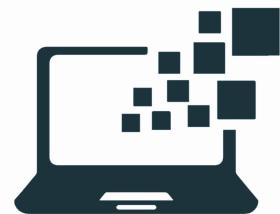
Examiner



Environment



Test



Technology

Key Questions: What platforms are there?

Wright et al (2020) suggests using only platforms set up for teleassessment (e.g., Presence Learning). Functionality, camera options, etcetera may introduce unknown variation. Others (e.g., C. Reynolds on listservs) has been more permissive, suggesting that secure versions of Zoom and other such teletherapy software are appropriate. The short answer, there's just too much unknown.

Key Questions: Can we work with local universities to establish standard error of measurement and other basic psychometrics?

Yes, this may help. However, it's a complicated discussion that effectively gets into study design. An informal approach wouldn't tell us anything we don't already know.

Key Questions: Can we work with local universities to establish a virtual normative sample.

Maybe? The short answer is that process is extremely technical. Even if you collected data, you'd be limited by design and sample issues. It's a much more complicated discussion than I want to get into here.

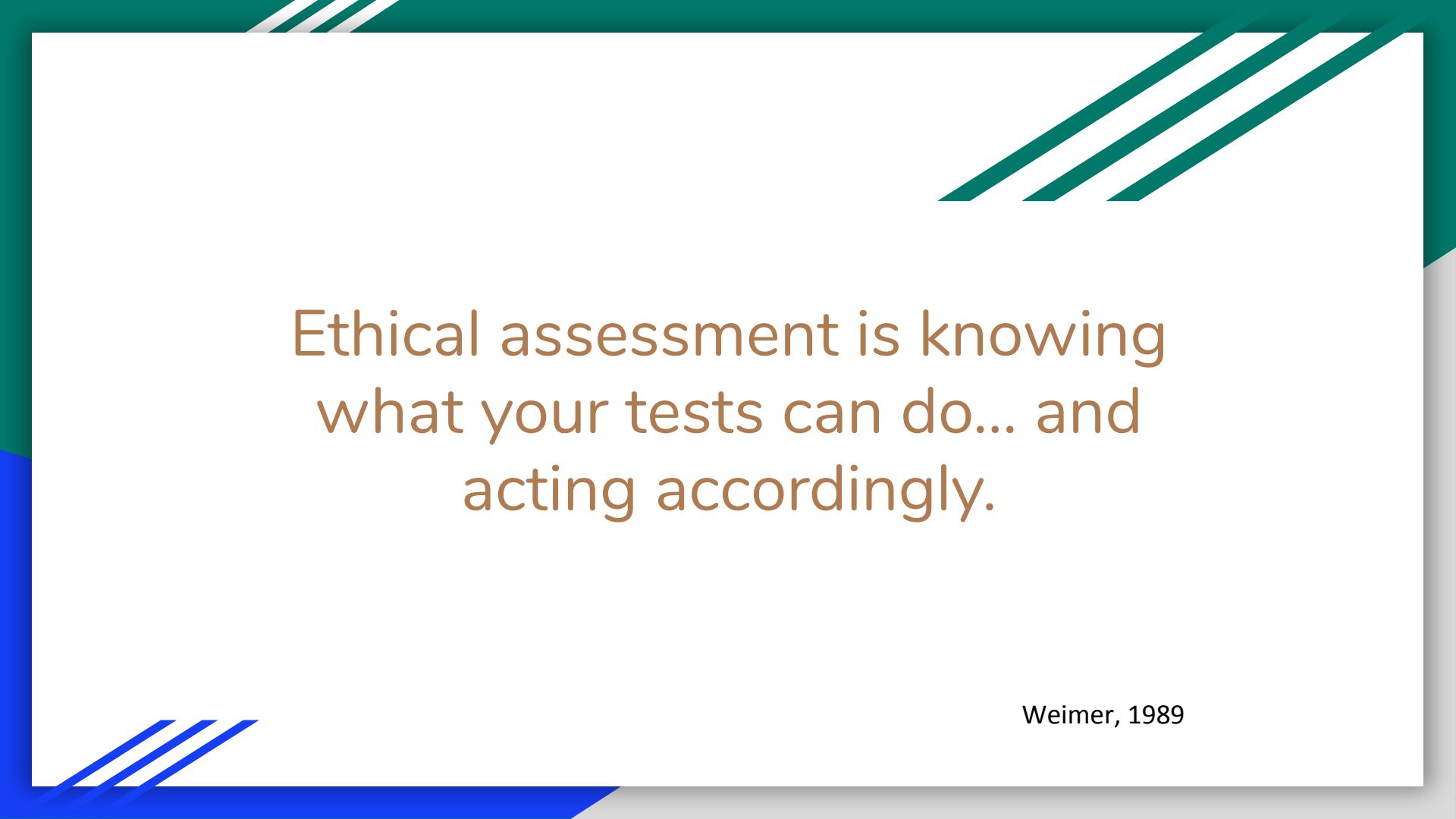
Key Questions: So what can we do?

I'm so glad you asked!

- Only give a standardized test when it's absolutely unavoidable and 100% necessary.
- When you give a test remotely, screen for unique sources of variance (e.g., STA, technology checks)
- When you are interpreting scores from a remotely administered test, be extremely conservative!
- Think carefully about the decisions your test scores impacted (e.g., classification, treatment) and progress monitor, progress monitor, progress monitor. It'll be on us to document efficacy of our decisions!

What does conservative interpretation look like?

- Avoid questionable interpretation strategies (e.g., profile analysis)
- Minimize use of less-reliable scores (they're likely even less reliable right now)
- Avoid using scores with inadequate validity evidence
 - Yes, this is generally always true. However, many practitioners get away with "in my experience..." to justify a score's use. In this case however, you don't have any experience with these scores (in this format). Avoid them!
- Use a 99% confidence interval! This will decrease your precision, but increase the likelihood that the true score is captured by your range.
- Corroborate. Test limits. Basically, look for evidence that your scores are wrong.



Ethical assessment is knowing
what your tests can do... and
acting accordingly.

Weimer, 1989

“It’s all about...first times and how hard it is to be new at things - from small things to global pandemics. When we have no relevant experience or expertise, the vulnerability, uncertainty, and fear of these firsts can be overwhelming. Yet, showing up and pushing ourselves past the awkward, learner stage is how we get braver.”

Brene Brown

How Do We Do This?



OCR Guidance

“If an evaluation of a student with a disability requires a face-to-face assessment or observation, the evaluation would need to be delayed until school reopens. Evaluations and reevaluations that do not require face-to-face assessments or observations may take place while schools are closed, so long as a student’s parent or legal guardian consents.”

Additional Guidance from NASP

“Some assessment publishing companies allow qualified examiners (school psychologists, school psychometrists and/or speech language pathologists) to use any test the qualified examiner deems appropriate even if the publishers have not intended for the test to be virtually administered. This raises issues regarding the reliability and validity of the results obtained, as norming samples have not included telehealth administration. Additional factors will need to be considered to maintain test integrity, including others who may be in the testing room and the level of distractions present during administration”.

Guiding Questions for School Teams



The following questions are not intended to contradict or supersede rules being set by district administration, but are offered as a resource when determining how to best move forward. For each of the following, providers are encouraged to consult with the child's IEP team and administration.

Guiding Questions: Evaluations

- What is best for the student?
- Does the team have means to access adequate data to complete a sufficiently comprehensive evaluation in order to determine eligibility?
- Is the parent or adult student comfortable proceeding with an evaluation during this time?
- Can the evaluation tools and tests be administered as standardized?
- Will modifying test materials and procedures to achieve physical distancing jeopardize test security?
- Have considerations of equity been discussed?

Guiding Question: Eligibility Decision-Making

- Is the evaluation sufficiently comprehensive?
- Is the student's performance due to a lack of appropriate instruction?
- Is the student's performance due to environmental or socioeconomic factors?
- Do the results reflect the student's functioning and needs in the school setting? Or are the results impacted by psychological, economic, and other stressors associated with the pandemic?
- Is appropriate caution being used when interpreting results?

Concerns and Considerations:

- A comprehensive and individualized evaluation requires direct classroom-based observation of the student in his or her natural learning environment.
- Standardized assessments (i.e., IQ tests, formal achievement tests) are not designed to be administered virtually and such tests must only be given through the means in which they were developed.
- If evaluations are to be conducted remotely, they should be conducted through platforms specifically designed for that purpose.

Concerns and Considerations:

- Providers across our state serve a diverse community of learners, many of whom are English language learners.
- Current circumstances may lead to validity issues which must be addressed when assessments are taking place in a time of heightened anxiety and uncertainty for youth, their families and caregivers, and school personnel.
- It is essential that legal and ethical guidelines related to student privacy be followed as methods for service delivery and evaluation move to virtual platforms.

- The mental health of students and staff is of utmost importance.
- All children should receive the academic, social, and emotional supports needed to be successful in the classroom.
- Before any evaluative tests are administered, make a sustained effort to establish a sense of calm and trust so that learning, and assessment of learning, can occur.





Considering Outside Evaluations During COVID-19



Are Private Psychologists Conducting Evaluations?

Many psychologists have paused psychological assessment services

Standardized administration procedures require in-person contact

- Standardized assessments often require manipulation of physical materials, standardized in-person interactions, and observations in a physical environment
- Virtual workarounds are likely to jeopardize test security and data quality/validity
- Research and evidence for equivalence of cognitive, neuropsychological, and other interactional measures is in the early stages and extremely limited



Are Private Psychologists Conducting Evaluations?

There is some leeway for psychologists to decide whether to proceed with modified virtual assessment procedures, use alternative measures available, or wait until in-person services are again feasible

- It is important for psychologists to know the limits of tele-testing and to consider if this approach is appropriate given the referral question, evidence, client characteristics/preferences and clinician expertise
- All psychologists are required to adhere to respective federal and state rules and regulations, the APA's Ethical Guidelines, and the Standards for Educational and Psychological Testing



Are Private Psychologists Conducting Evaluations?

Some private psychologists are resuming in-person face-to-face assessments. **Things to consider:**

Impact of PPE and any other deviations from standardized administration

- Has the psychologist described any social distancing measures and considered their impact on validity?
- If this information is not documented, it would be reasonable to call the evaluator for clarification

ADOS-2 guidance from developer, Dr. Cathy Lord

- Cannot be administered and **scored** with mask(s) or use of other PPE
- Stop-gap tool in development - called the Brief Observation of Symptoms of Autism (BOSA) - FREE
 - Only to be administered by professionals trained and competent in ADOS-2 administration
 - Requires ADOS-2 protocols and kit, with some additional items/toys needed
 - More information about upcoming trainings here:
<https://www.semel.ucla.edu/autism/event/brief-observation-symptoms-autism-bosa-trainings>

Outside Evaluation Data

Recommended Steps for Reviewing Outside Evaluation Data

- Provide a copy of the report to the qualified examiner(s) at the school
- The school's qualified examiner(s) should enter scored data into Review of Existing Data
- The school team will review the new information as it relates to existing school-based information and to suspected eligibility category

Qualified Examiner



Multidisciplinary Team

Outside Evaluation Data



Recommended Steps for Reviewing Outside Evaluation Data

- For any required evaluation components not included as part of the outside evaluation and not already documented in the student's records, the team will need to obtain additional data
- For any required evaluation components that do not match the current school functioning and existing data, the school will need to consider if additional data are necessary to confirm functioning
- If all required evaluation components are included in the outside evaluation and the data are consistent with the student's school functioning, then the team may proceed without collecting additional data



Thank
you!!

References

- Callan Stoiber, K. (2014). 3: Problem solving foundations for school psychological services. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 25-40). Bethesda, MD, MD: NASP.
- Callan Stoiber, K. (2014). 3: A comprehensive framework for multitiered systems of support in school psychology. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Data-based and collaborative decision making* (pp. 41-70). Bethesda, MD, MD: NASP.
- Daniel, M., & Wahlstrom, D. (2019). Raw-score equivalence of computer-assisted and paper versions of the WISC-V. *Psychological Services, 16*(2), 213-220. <https://doi.org/10.1037/ser0000295>
- Farmer, R., Lockwood, A., Harris, B., McClain, M. B., Dombrowski, S. C., & McGill, R. J.. (2020). Questions to Ask Before Remote testing (Version 1). figshare. <https://doi.org/10.6084/m9.figshare.12330533.v1> ([<https://osf.io/axp5s/>])
- Farmer, R. L., McGill, R. J., Dombrowski, S. C., McClain, M. B., Harris, B., Lockwood, A. B., Powell, S. L., Pynn, C., Smith-Kellen, S., Loethen, E., Benson, N. F., & Stinnett, T. A. (in press). Teleassessment with children and adolescents during the Coronavirus (COVID-19) pandemic and beyond: Practice and policy implications. *Professional Psychology: Research and Practice*.
- Florida Department of Education. (2012). *Applying Florida's Planning and Problem-Solving Process (Using RtI Data) in Virtual Settings*. <http://www.fl DOE.org/core/fileparse.php/7509/urllt/0085404-mtss-virtual.pdf>

References

- Miciak, J., Fletcher, J. M., (2020). The critical role of instructional response for identifying dyslexia and other learning disabilities. *Journal of Learning Disabilities*. 1-11. <https://doi.org/10.1177%2F0022219420906801>
- National Association of School Psychologists. *Problem-solving teams: Information for parents and educators*.
https://www.nasponline.org/assets/documents/Resources%20and%20Publications/Handouts/Families%20and%20Educators/nasp_probsolve.pdf
- Wright, A. J., Mihura, J. L., Pade, H., & McCord, D. M. (2020, May 1). *Guidance on psychological tele-assessment during the COVID-19 crisis* [Brief]. Retrieved from <https://www.apaservices.org/practice/reimbursement/health-codes/testing/tele-assessment-covid-19>
- Weiner, I. B. (1989). On competence and ethicality in psychodiagnostic assessment. *Journal of Personality Assessment*, 53, 827-831. https://10.1207/s15327752jpa5304_18
- Wright, A.J. (2018a). Equivalence of remote, online administration and traditional, face-to-face administration of the Woodcock-Johnson IV cognitive and achievement tests. *Archives of Assessment Psychology*, 8(1), 23-35.
https://www.presencelearning.com/app/uploads/2016/09/WJ-IV_Online_Remote_whitepaper_FINAL.pdf
- Wright, A. J. (2018b). Equivalence of remote, online administration and traditional, face-to-face administration of the Reynolds Intellectual Assessment Scales-Second Edition. [White paper]. Psychological Assessment Resources, Inc.
<https://www.parinc.com/Products/Pkey/365>

QUESTIONS TO ASK BEFORE REMOTE TESTING



IS THE TECHNOLOGY AVAILABLE?

Do you and the student have access to the hardware and internet?

Teleassessment requires a lot of infrastructure. Both the examiner and the student need a computer, at least one (but preferably two) good cameras, headphones and microphone, and broadband internet.



WILL THE ENVIRONMENT WORK AGAINST ME?

Is the setting too loud or distracting?

Testing remotely has its perks. However, examiners won't be able to control the environment as much as they're used to (and no, the research doesn't account for this!). Distractions, interference from family members, and competing activities in the home should be considered.



ARE THOSE DATA ESSENTIAL?

Will the data be used to make time-sensitive decisions?

The utility of standardized data is limited to helping out during diagnosis/classification. Is that information you need right now? Are there alternatives? Would you have enough data to make a decision either way?



WHAT ABOUT CULTURE & LANGUAGE?

Is the test culturally and linguistically appropriate?

Assuming that test norms are representative of the student (you should check), do you need an interpreter to obtain consent or conduct interviews prior to assessment? Are your tele-consent forms translated? How about other materials?



WILL MY SCORES BE VALID?

Does the student's behavior, energy level, or emotions interfere with testing?

This is a stressful time! You'll need to check that you're getting the best performance from the student. This means screening for emotional concerns, watching for fatigue, asking about stressors ahead of time, and keeping an eye out for other irrelevant behaviors that could interfere with testing.

→ IF YOU CHOOSE TO TEST, DOCUMENT EVERY STEP AND INTERPRET THE RESULTS WITH EXTREME CAUTION!

