

Exploring Assessment in Professional Psychology

New Tests Are Coming: Now is A Good Time to Stop Making Mistakes in Your Administration and Scoring

By Michael Lee Zwiers, Ph.D., R. Psych

Sometime in 2024, we can expect the newest editions of the Wechsler Adult Intelligence Scale (WAIS-5) and the Stanford Binet Intelligence Scale (SBIS-6) to be released, followed by the 4th Edition of the Wechsler Memory Scales in late 2025. The Canadian editions will likely arrive within a year. We will have to learn the correct administration of these tests. Unfortunately, research dating back to the 1970s tells us that psychologists make errors when administering and scoring intelligence tests (>90% in a 1979 study). Students make an average of 10 to 15 scoring errors per protocol, and even students under close supervision make about 5 errors per protocol. Even after they get it right, drifting and errors begin to recur.

So when the new tests arrive, we can't assume that we just have to learn how to administer the new subtests and we will be ready to go. We should be reviewing the entire test in detail, (particularly to identify minor changes to administration or scoring so that we don't simply follow old habits). For those who are curious, the WAIS-5 has 3 new subtests (Set Relations, Naming Speed Quantity, and Running Digit Span). The WMS-5 has eliminated 8 subtests, has introduced 4 new subtests (Symbol Span, Spatial Addition, Design Memory and General Cognitive Screener), and made modifications to 3 subtests.

Unfortunately, the title of this article is not just provocative—it accurately depicts how the vast majority of psychologists conduct formal test administration. Yes, I mean you. None of us is perfect. And unfortunately, we make errors. Recent studies suggest that 94% to 98% of protocols contain at least one administration or scoring error. And that's based on only what can be gleaned from the protocols themselves. Some of these mistakes are administrative errors, others are judgment errors, while others are clerical errors (incorrect addition, wrong table, wrong column).

I have trained graduate students in psychology, trying to improve their test administration and scoring. This has included live modeling with descriptive advice; followed by close observation and feedback of live practice administration; followed by scripted one-to-one administration of the full test to the instructor or doctoral graduate student TA, including challenge of all prompts to ensure competence; after which they conduct videotaped administration with practice clients scored by the instructional team. Even with this wrap-around support and multilevel feedback, student administrators continued to make errors. And even when errors are reduced through careful instruction and immediate feedback, drifting occurs in later administrations post-evaluation.

What should we do about this? In the rest of this article, I will provide some suggestions to help improve your administration and scoring so that you can stop making mistakes. And catch and correct the ones you do make.

- 1) **“Perfect Practice Makes Perfect.”** When the new tests arrive, go through every detail of the test administration. Read the manual carefully and highlight administration, feedback, and scoring guidelines. Then, sit down with a colleague

and practice administering the test without making mistakes. Ideally, both examiner and examinee will have a copy of the manual. Give each other feedback. If someone does make a mistake, fix it immediately and over-practice the correct administration to ensure you learn it correctly.

- 2) **Audit your current practices.** Even before those new tests arrive, you should sit down with your manual(s) and go over every step of every subtest, including the general administration and scoring procedures. **If you want to improve your practice, first assume that you are doing something wrong. Then make it your goal to find that error so that you can correct it.** An easy way to do this is to have a colleague observe your administration at least once a year. Or meet with a colleague to swap administrations with each other.
- 3) **The most common administration errors include not querying properly, not adhering to time limits, not accurately recording completion times, and neglecting to provide teaching feedback** when required on early test items where participants do not earn full points. Note: this feedback is critical to ensure that participants have the opportunity to better understand test requirements and improve their subsequent responses. If you use digital administration, you need to click the *Teaching* button or the *Query* button. If hand administering, you should record a (T) or (Q). To time properly, you need to know when precisely to start timing (hint: timing typically starts as soon as you finish reading the oral instructions, which are typically repeated every time and not abbreviated). The “*Do you have an answer?*” query does not need to be asked at full volume but does need to be pronounced when 10 seconds remain.
- 4) **The most common scoring errors are related to verbal items (Similarities, Vocabulary, Information, Comprehension).** When to query? When not to query? Basically, you should only query when the response provided by the examinee matches a response with a **(Q)** in the manual. Or, in rare circumstances when you are confused about what the person is saying. A simple way to reduce query errors is to memorize the most common 2-point answers. Next, review some of the 0-point answers (these are usually easy to identify). If the participant response doesn't match a 0-point or 2-point response then review the 1-point answers during the testing session to see if the response matches. If it doesn't, quickly scan all of the **(Q)** responses to make sure none were missed. That way, you can clarify your rating later on and be relatively confident that you haven't missed a requisite query. When scoring later on, be sure to read the item-scoring parameters at the top of each question. This provides a broad framework of understanding for what qualifies as a 2-point or a 1-point question. My final suggestion is to consult with a colleague for any items that you are not sure about. This will help both of you to become better at scoring.
- 5) **Keys to a few subtests:** Arithmetic. If an examinee requests you to repeat a question, you have to pause your stopwatch until you finish re-reading it. One-per-Second Timing: the most common error on Digit Span or Letter-Number Sequencing is administering items either too fast or too slow. I keep an analog clock in the testing room that ticks audibly every second. Keeping time then becomes easy.

- 6) **Read the instructions from the script. Every time.** Think you can memorize the script? Think again. Even professional actors will make mistakes on stage with well-rehearsed material.
- 7) **Ceiling rules vary.** Review each subtest in your record book before administering the subtest.
- 8) **All hand scoring needs to be re-checked.** This includes rechecking the age of the participant. Why? Because **humans are prone to making errors, and we need to catch those errors.** Pharmacists use a 3-time check system to ensure the right person gets the right medication in the right dose. I might remind you that surgeons have only recently started using checklists to ensure that they count every sponge, piece of gauze, or clamp that goes into someone so they can be sure it is removed at the end of their surgery. Surgical checklists were instituted because too many unnecessary medical errors were being made. When adding scores, I suggest checking the math twice more by doing the calculations in 3 different ways. For example, add from top to bottom, then recheck by adding from bottom to top; add the numbers in sequence, then recheck by adding the 2's followed by the 1's; begin with your ceiling item and subtract all of the zeros. When re-checking scoring tables, use the standard score to track back to the raw score range.
- 9) If you are looking for a detailed test administration checklist, purchase one of Jerome Sattler's textbooks. He provides crystal clear guidance on administration and scoring for every single subtest, including special queries, clarifications, or responses. Hopefully, his publishing company will continue to include this information after he fully retires and passes along the reins.

Whether you're using familiar tests or learning new ones, be sure that you are not one of those psychologists who makes administration or scoring errors.

Dr. Michael Lee Zwiers R. Psych.