**If your goal is to explain—credibly and in plain English—why high-stakes standardized reading tests can undervalue real, documented gains for a student with profound dyslexia (even one making clear progress deep into Wilson Step \_\_\_\_), these Ph.D.-level researchers are top fits to testify:**

1. **Hugh W. Catts, Ph.D.** (Florida State University; former Univ. of Kansas)  
   Best for: Explaining why state reading tests are **poor benchmarks** for individual student growth and what to use instead. He’s written directly on the misalignment between high-stakes tests and meaningful reading outcomes. [The Reading League+1](https://www.thereadingleague.org/wp-content/uploads/2022/02/JanFeb2022-TRLJ-Article.pdf?utm_source=chatgpt.com)
2. **Jack M. Fletcher, Ph.D.** (University of Houston; NICHD-funded, past president of the International Neuropsychological Society)  
   Best for: SLD identification, RTI/MTSS, and assessment validity. He can connect IDEA’s “multiple measures” requirement to why a **body of evidence** (program mastery checks, lesson completion, progress-monitoring) can outweigh a single standardized score. [NIH News in Health+2Wyoming Instructional Network+2](https://newsinhealth.nih.gov/2024/02/dr-jack-fletcher-dyslexia?utm_source=chatgpt.com)
3. **Richard K. Wagner, Ph.D.** (Florida State University; Florida Center for Reading Research; co-author of CTOPP-2, TOWRE-2)  
   Best for: What common standardized tools actually measure (speed/efficiency) vs. what your Wilson data show (accurate decoding mastery over time). He’s ideally positioned to testify to test **limitations and proper interpretation**. [fcrr.org+2International Dyslexia Association+2](https://fcrr.org/people/richard-wagner-phd?utm_source=chatgpt.com)
4. **Guinevere Eden, Ph.D.** (Georgetown University; past IDA presidential awardee)  
   Best for: Neuroscience of dyslexia—why students can improve substantially with explicit instruction yet still underperform on certain timed, print-heavy measures due to the nature of the disability, not lack of learning. She has testified to the U.S. Senate on dyslexia science. [U.S. Senate HELP Committee+2APM Reports+2](https://www.help.senate.gov/imo/media/doc/Eden.pdf?utm_source=chatgpt.com)
5. **Maryanne Wolf, Ed.D.** or **Virginia W. Berninger, Ph.D.** (emerita scholars) — strong complementary voices if you need writing/spelling and language components tied in; Berninger has long argued for instruction-aligned measures and growth data. [UW College of Education+1](https://education.uw.edu/news/feature/study-shows-specific-learning-disabilities-differ-unit-language-affected?utm_source=chatgpt.com)

Why these picks map to your exact point

* **High-stakes standardized reading tests = weak growth monitors** for dyslexia, especially when timed and comprehension-heavy; they often lag behind the real improvements captured in structured-literacy progress data (word reading/decoding accuracy, phonemic proficiency, nonsense-word reading, mastery checks). Catts lays this out explicitly. [The Reading League](https://www.thereadingleague.org/wp-content/uploads/2022/02/JanFeb2022-TRLJ-Article.pdf?utm_source=chatgpt.com)
* **IDEA requires multiple measures**; relying on a single standardized score to deny or discount programming is unsound. Fletcher’s RTI/Hybrid-Model work is the go-to here. [Wyoming Instructional Network](https://wyominginstructionalnetwork.com/wp-content/uploads/2018/05/Identifying-Learning-Disabilities-in-the-Context-of-Response-to-Intervention_A-Hybrid-Model.pdf?utm_source=chatgpt.com)
* **Test construction matters**: Wagner can explain how speeded measures (e.g., efficiency/fluency) can remain depressed longer—even when accuracy and decoding skills have **truly improved**—which is typical in profound dyslexia. [International Dyslexia Association+1](https://dyslexiaida.org/why-is-it-so-difficult-to-diagnose-dyslexia-and-how-can-we-do-it-better/?utm_source=chatgpt.com)
* **Neuroscience evidence** shows dyslexic readers can learn to read significantly better with explicit instruction while still showing group-norm deficits on certain tasks; Eden can translate that science clearly for a fact-finder. [U.S. Senate HELP Committee+1](https://www.help.senate.gov/imo/media/doc/Eden.pdf?utm_source=chatgpt.com)

How to brief your expert (so they can say this cleanly)

* Provide:  
  • Wilson lesson history and **mastery checks** (through Step \_\_\_\_\_\_), session fidelity logs, and error-analysis trends.  
  • Curriculum-based measures: **nonsense-word reading**, **decoding accuracy**, **phoneme segmentation/blending** growth, controlled-text passage accuracy.  
  • Any general-ed/state test results—so the expert can contrast **what they measure** vs **what Wilson progress demonstrates** (skill acquisition that standardized tests under-sample).  
  • IDEA citations on “variety of assessment tools and strategies” to reinforce the multiple-measures argument (your legal handouts already hit this). (Source note for you: 20 U.S.C. §1414; 34 C.F.R. §300.304.)

Practical next steps

* Start with **Catts** or **Fletcher** as primary; add **Wagner** for testing-interpretation depth and **Eden** if you want a neuroscience lens.
* Reach out via their university/lab pages and ask about **expert testimony/consulting** availability (many academics do this selectively). Pages to start: Catts and Wagner (FSU/FCRR), Fletcher (Univ. of Houston), Eden (Georgetown). [U.S. Senate HELP Committee+3fcrr.org+3ResearchGate+3](https://fcrr.org/people/richard-wagner-phd?utm_source=chatgpt.com)