

Testimony of Matt DeCarlo, MSW, Ph.D.
February 4, 2025
Maryland Senate Finance Committee
In Support of Senate Bill 0379

Good Afternoon Chair Beidle, Vice Chair Hayes, and distinguished members of the Senate Finance Committee,

My name is Matt DeCarlo and I am the MSW Program Director at Saint Joseph’s University and co-founder of Open Social Work, which advances open practices in our discipline. I am testifying in my personal capacity as a social worker (not on behalf of my employer) and professional capacity as a social work researcher and advocate.

My research focuses in large part on open workforce reforms, and specifically, social work licensure. I appreciate the opportunity to offer testimony in support of SB0379.

Please see the attached fact sheets, grounded in my years of research on the Association of Social Work Boards examination program.

ASWB: A Uniquely Profitable “nonprofit” 2

Recession-Proof Profits.....2

How Profitable is ASWB?..... 2

Educated but excluded.....3

ASWB Secretly Deletes Biased Items..... 4

ASWB Monitors & Removes Scored Items for DIF..... 5

Invalid Licensure Decisions.....6

A Growing Problem..... 6

ASWB Violates Psychometric Standards..... 7

New ASWB Questions are 8% Easier..... 7

Are 2018 Blueprint Cut Scores Still Valid?.....8

ASWB Ignores Required Validation & Equity Tests..... 9

Validation is Key for Exam Precision..... 10

ASWB Exams Are Inadequately Documented..... 11

ASWB Clearly Violates Documentation Standards..... 12

ASWB Ruined Its Exams’ Reliability..... 13

Broken Exams, Broken Workforce..... 14

White Test-Takers Improve, Others Worsen..... 14

Exam Pass Rates = Licensure Rates..... 16

When Illinois Eliminated the Exam..... 16

ASWB: A Uniquely Profitable “nonprofit”

In 2010, [Albright & Thyer](#) found that MSW students could easily guess the correct answer on Association of Social Work Boards (ASWB) practice examinations without knowing the question, simply based on how answers were phrased. In their abstract, they stated:

“The multimillion dollar social work testing industry is big business; both the profession and the public, which relies heavily on the gate-keeping function of these tests, deserve greater transparency and accountability with respect to their legitimacy” (p. 229)

ASWB has not provided any transparency on the validity or reliability of its exams. However, ASWB is required to be transparent in tax reporting.

Using the [ProPublica Nonprofit explorer for ASWB](#), this brief collects data reported by ASWB to the Internal Revenue Service (IRS). It also uses ASWB’s annual reports posted to the [aswbannualreport.org](#) website from 2015-2021. Finally, it uses data from [ASWB’s Exam Pass Rate Analysis](#), which differentiates between first-time and repeat test-takers from 2011-2021.

Recession-Proof Profits

Since 2011, ASWB has posted recession-proof profit margins rivaling Fortune 500 companies.

- From 2011 to 2021, ASWB’s net assets increased by 4.5x.
- ASWB’s net assets in 2021 were over \$40,000,000, up from \$9,000,000 in 2011.
- In 2021, ASWB’s profit margin was 29% with \$24,000,000 in total revenue.
- ASWB’s profit margin has averaged 17% since 2011.
- ASWB holds a \$30 million investment portfolio managed by Morgan Stanley.

How Profitable is ASWB?

	2011	2012	2013	2014	2015	2016
LCSW first-time test-takers	9,100	9,604	10,879	12,217	13,044	14,007
LMSW first-time test-takers	11,260	12,732	13,110	14,184	15,214	15,496
LBSW first-time test-takers	3,164	3,251	3,595	3,873	4,083	4,113
Repeat test-takers	--	--	--	--	11,457	11,127
ASWB exam revenue	--	--	--	--	\$ 11,345,503	\$ 11,689,451
ASWB publishing revenue	--	--	--	--	\$ 1,686,190	\$ 1,489,956
ASWB net assets	\$ 8,995,137	\$ 10,612,898	\$ 13,079,412	\$ 15,680,159	\$ 17,750,104	\$ 19,693,116
ASWB total revenue	\$ 9,461,425	\$ 10,279,908	\$ 11,492,614	\$ 12,692,553	\$ 13,964,190	\$ 13,767,709
ASWB profit margin	10.61%	14.57%	20.64%	22.26%	19.47%	8.45%

	2016	2017	2018	2019	2020	2021
LCSW first-time test-takers	14,007	16,095	16,022	17,207	16,801	20,657
LMSW first-time test-takers	15,496	16,884	16,812	18,231	16,716	21,650
LBSW first-time test-takers	4,113	4,462	3,711	3,583	2,709	3,494
Repeat test-takers	11,127	12,617	13,478	19,526	15,521	20,720
ASWB exam revenue	\$ 11,689,451	\$ 12,996,484	\$ 13,277,694	\$ 14,410,319	\$ 13,735,930	\$ 17,659,247
ASWB publishing revenue	\$ 1,489,956	\$ 1,562,915	\$ 1,704,718	\$ 1,751,079	\$ 1,684,234	\$ 2,123,620
ASWB net assets	\$ 19,693,116	\$ 23,289,150	\$ 24,046,614	\$ 28,831,413	\$ 33,841,553	\$ 40,273,169
ASWB total revenue	\$ 13,767,709	\$ 15,565,636	\$ 16,344,808	\$ 17,595,886	\$ 16,234,758	\$ 24,599,963
ASWB profit margin	8.45%	14.85%	16.52%	15.56%	18.46%	29.08%

Educated but excluded

Exams are a bottleneck for aspiring social workers, and the problem is getting worse.

- Repeat test-takers comprise about 30% of all LCSW, MSW, and BSW examinees
- Since 2011, the number of repeat test-takers grew by 86%.
- Repeat-test takers grew at twice the rate as first-time examinees (43%).
- In 2021 and 2019, there were more repeat test takers than first-time LMSW examinees.

Repeat test-takers calculation: I subtracted first-time test-takers reported in ASWB's Exam Pass rate Analysis from the total test takers reported in ASWB's Annual Report. 2021 was the only year that ASWB reported first-time test-takers of Advanced Generalist and Associate levels of exams, so I applied those figures to previous years as a best guess. There are few test-takers at these levels. All other data are copied directly from ASWB's public reports.

ASWB Secretly Deletes Biased Items

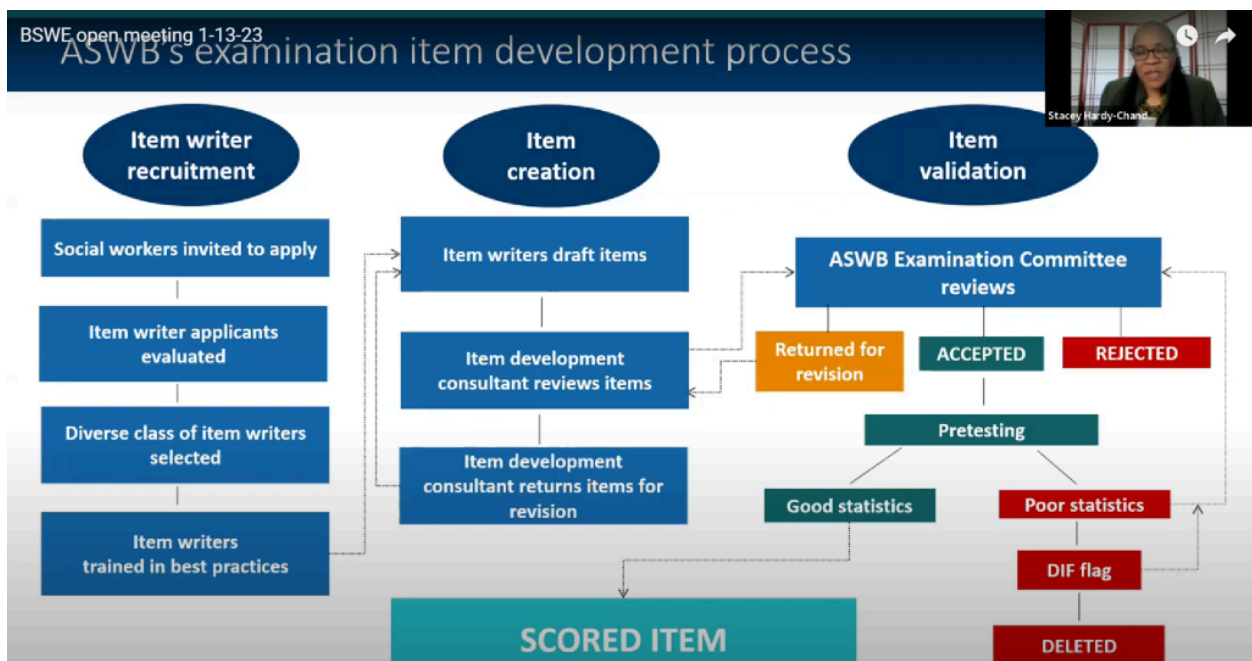
ASWB follows psychometric best practices by pretesting all items for Differential Item Functioning (DIF) before they are used as scored items on the examination. Items that demonstrate DIF are removed or revised by subject matter experts. Revised items are pretested again to demonstrate that DIF has been resolved prior to moving items into the live, scored section of the examination. As ASWB's CEO, Dr. Stacey Hardy Chandler stated [in a zoom webinar with Dr. Jennifer Klafehn](#), a representative from ASWB's psychometrics consultant:

SHC: I actually want to start and clarify because if I were in the audience, I would want clarity on this. You talked about items being flagged for DIF. Can you clarify at what point those items are deleted. In other words, the test item phase in terms of impacting scores.

JK: So they are deleted at the pretesting phase.

SHC: Just wanted to clarify that those do not impact scores when they are identified. (ASWB, 2023b, 53:39)

This is mirrored in the flowchart ASWB provides to its member boards about the DIF procedure.



According to these ASWB sources, scored items are *never* removed for statistical bias.

ASWB Monitors & Removes Scored Items for DIF

Contrary to these statements, ASWB's public statements (a small sample of which are reproduced below) describe the procedure for evaluating *and removing scored items* for bias.

In the [current ASWB examination guidebook](#) ASWB explains:

“psychometric experts use statistical tracking of responses to determine whether each question measures competence effectively and fairly. Exams are built using only scored questions that have met these rigorous standards. **All questions continue to be monitored by psychometric experts to prevent any scored questions from being used that do not meet standards for fairness.**” (p. 8)

“All ASWB examination questions are monitored to ensure that there is nothing in the content that would provide an advantage to one demographic group over another. This evaluation occurs during the pretest phase—before questions are included in the scored question bank—and **continues while they are used as scored questions**” (p. 21).

In a [2021 issue of the New Social Worker](#), Stacey Owens, ASWB Item Development Consultant gave rough statistics:

The first tool ASWB uses is a testing industry statistical measurement called Differential Item Functioning (DIF). DIF indicates whether an exam question shows tendencies to advantage or disadvantage one group of test-takers over another (ASWB, 2020). DIF is identified by statistically analyzing responses to the exam questions—called items—during pretesting. **Scored items are continually monitored for DIF.** On an annual basis, less than 5% of all items released show DIF. Items flagged for DIF are removed from the bank of potential exam questions. (para. 3)

In a [2020 article in Social Work Today](#), ASWB Examination Director Lavina Harless confirmed:

“Monitoring of item performance doesn't end once an item moves out of pretest status. **Scored items are continually monitored** to ensure that performance doesn't slip. **If a scored item demonstrates a statistically significant drop** in performance, it is **taken out of use and returned to the examination committee for review.** Should the committee decide to edit and keep the item, it returns to pretest status” (para. 12).

In ASWB's [2023 Testing Enhancements webinar](#), Examination Director Harless detailed:

“What if then people start to get it wrong? Again, because we do our research ahead of time, that doesn't typically happen. **If we see something that looks like maybe there's a shift in practice or something occurring, we'll pull those test questions down** and take a look at those with our subject matter experts on our Examination Committee” (57:20)

Invalid Licensure Decisions

It is common for social workers to fail the ASWB examination by 1 or 2 points. If 1 or 2 items were later found by ASWB to have bias, it would invalidate the licensure decision for that person.

Test-takers and boards have a right to know when a scored exam item, used to make a licensure decision, was later removed by ASWB because it demonstrated statistical bias. For this reason, ASWB exams are a poor fit for licensure because the true cut score is likely to change over time.

If ASWB followed the code of ethics and reported when exam items removed for bias impacted licensure decisions, social work boards would find it difficult to license social workers in a predictable and legally defensible manner.

A Growing Problem

ASWB publicly reported 5x higher DIF in its 2021 exams than in 2010 exams.

- In 2011, [Marson and colleagues](#) reported less than 1% of exam items showed DIF.
- In 2021, [Stacey Owens](#) reported less than 5% of items were flagged for DIF.

ASWB examinations are 170 questions long, with 20 unscored items.

- DIF appears in up to 8 or 9 questions (5%) of the 170 question exam.

ASWB reports this “usually” does not happen, but that when it does, there is no procedure for notifying the affected parties. This is a grievous ethical lapse.

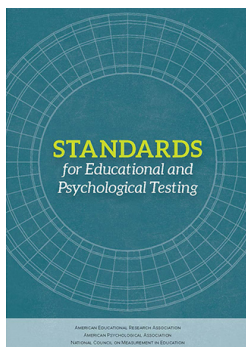
[ASWB's \(2021\) Manual for New Board Members](#) states “when consistent DIF is identified in an item—**usually** [emphasis added] in pretest items that are being pretested for possible scored use—the item is deleted from the item bank” (p. 20).

In their blog post on measurement fairness and [accompanying PowToons video](#) ASWB states that “when an item shows DIF, **usually** [emphasis added] in the pretest stage, the question goes no further” (1:30).

How many scored items has ASWB removed for biased statistical functioning? They won't share data.

ASWB Violates Psychometric Standards

ASWB cites as its principal source of psychometrics the [Joint Standards for Educational and Psychological Testing](#) by the National Council on Measurement in Education, the American Educational Research Association, and the American Psychological Association.



“The purpose of the *Standards* is to provide criteria for the development and evaluation of tests and testing practices and to provide guidelines for assessing the validity of interpretations of test scores for the intended test uses. Although such evaluations should depend heavily on professional judgment, the *Standards* provides a frame of reference to ensure that relevant issues are addressed. All professional test developers, sponsors, publishers, and users should make reasonable efforts to satisfy and follow the Standards and should encourage others to do so. **All applicable standards should be met by all tests and in all test uses unless a sound professional reason is available to show**” (p. 1)

“Although the *Standards* is not enforceable by the sponsoring organizations, it has been repeatedly recognized by regulatory authorities and courts as setting forth the generally accepted professional standards that developers and users of tests and other selection procedures follow. Compliance or noncompliance with **the Standards may be used as relevant evidence of legal liability in judicial and regulatory proceedings**. The Standards therefore merits careful consideration by all participants in the testing process” (p. 2)

ASWB follows some of these standards, but it ignores important standards during the exam validation process. Because of these violations, exam cut scores are systematically biased.

New ASWB Questions are 8% Easier

ASWB announced [Testing Experience Enhancements on their blog](#). Beginning in January 2023, ASWB changed their examination format from 4-option questions (A, B, C, or D) to 3-option questions (A, B, or C) “with the goal of completing the transition by 2025” (para. 3).

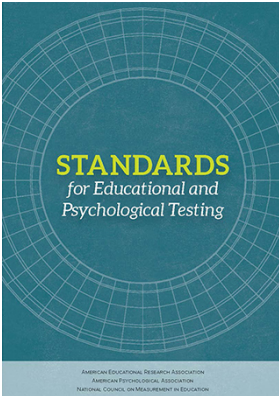
It is true that the choice between 3-option and 4-option questions can be equally valid, but it is statistically unknown what the impact on the validity of cut scores would be.

- Items without a 4th option are 8% easier to guess correctly (25% vs. 33%)
- 3 options take 25% less time to read than 4 options across a 4-hour, 170-item exam.

If the exam is easier, the cut score should increase.

Are 2018 Blueprint Cut Scores Still Valid?

Changing the number of answer options is a major, untested, shift. ASWB did not calculate new cut scores for a mix of 3 & 4-option exams. This is a violation of the *Standards* which require:



Standard 7.14: When substantial changes are made to a test, **the test's documentation should be amended, supplemented, or revised** to keep information for users current and to provide useful additional information or cautions.

Comment: ...**When substantial changes are made to items and scoring, information on the extent to which the old scores and new scores are interchangeable should be included in the test documentation.** Sometimes it is necessary to change a test or testing procedure to remove construct-irrelevant variance that may arise due to the characteristics of an individual that are unrelated to the construct being measured (e.g., when testing individuals with disabilities). **When a test or testing procedures are altered, the documentation for the test should include a discussion of how the alteration may affect the validity and comparability of the test scores, and evidence should be provided to demonstrate the effect of the alteration on the scores** obtained from the altered test or testing procedures, if sample size permits.

Clearly, ASWB makes changes to their examinations without proper psychometric evidence showing the altered test format is psychometrically equivalent to the old examinations.

ASWB is required to test whether these three groups have equivalent scores. They did not do so.

1. Pre-2023 Exams with 170 4-option questions.
2. 2022-2025 Exams with an unknown mix of 3-option and 4-option questions.
3. 2025 Exams and beyond, with 3-option questions.

ASWB announced that eliminating an answer option would address concerns about construct-irrelevant variance—“reducing time pressure and ensur[e] a focus on a test-taker’s social work knowledge” (para. 3). Clearly ASWB hoped that using 3-option items would improve test scores.

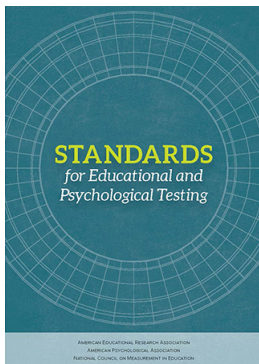
But it never re-calculated cut scores with the new question format. Is it true, as ASWB says in their Exam Guidebook, that “no test-taker receives an advantage or disadvantage because of the version of exam received (p. 20)?

ASWB reports never testing if 3- or 4-option exams are psychometrically equivalent, but ASWB’s examination program relies on the equivalency of exams across years.

ASWB Ignores Required Validation & Equity Tests

ASWB started using this new question format in the wake of the 2022 Exam Pass Rate Analysis which demonstrated systematic bias in favor of younger white social workers.

Contrary to ASWB's response to exam inequities, the *Standards* do not direct test-makers like ASWB to eliminate an answer option when systematic biases emerge. Rather, in Standard 2.15, the *Standards* require ASWB to conduct at least one of two specific psychometric tests to estimate the multivariate impact of race, age, and language on scores.

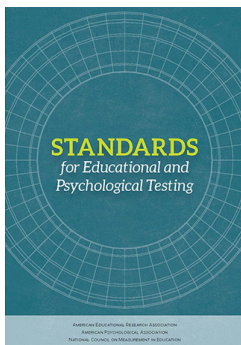


Standard 2.15: When there is credible evidence for expecting that **conditional standard errors of measurement or test information functions** will differ substantially for various subgroups, investigation of **the extent and impact of such differences should be undertaken and reported as soon as is feasible.** (emphasis added)

Comment: If differences are found, they should be clearly indicated in the appropriate documentation. In addition, if substantial differences do exist, **the test content and scoring models should be examined** to see if there are legally acceptable alternatives that do not result in such differences. (p. 46)

ASWB does not calculate the Conditional Standard Error of Measurement or Test Information Function for minoritized social workers. Why is this important? The *Standards* are clear:

“When a test score or composite score is used to make classification decisions (e.g., pass/fail, achievement levels), the standard error of measurement at or near the cut scores has important implications for the trustworthiness of these decisions...



Standard 2.14: When possible and appropriate, conditional standard errors of measurement should be reported at several score levels unless there is evidence that the standard error is constant across score levels. **Where cut scores are specified for selection or classification, the standard errors of measurement should be reported in the vicinity of each cut score.**

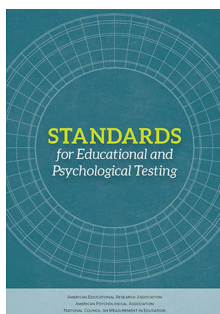
Comment: Estimation of conditional standard errors is usually feasible with the sample sizes that are used for analyses of reliability/precision. **If it is assumed that the standard error is constant over a broad range of score levels, the rationale for this assumption should be presented.** The model on which the computation of the conditional standard errors is based should be specified” (p. 47).

ASWB does not calculate or report conditional standard errors or test information functions of its exams, despite the clear mandate from Standard 2.15 to do so *as soon as is feasible*.

Validation is Key for Exam Precision

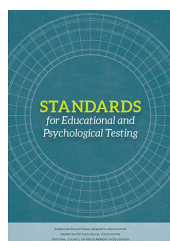
ASWB examinations use Item Response Theory (IRT) to evaluate exam items, but they do not use it to evaluate the exam as a whole. [Bean \(2022\)](#) highlights how IRT tests like those required in the *Standards*—conditional standard error of measurement or test information function—reveal important information for validity, reliability, and fairness.

In common terms, IRT analyzes whether the examination is appropriately precise *for test-takers whose true ability is near the cut score*. An examination for licensure should be most precise at the decision-point. An examination that was most precise for excellent or abysmal social workers would be invalid for use as an examination. For this reason, calculating the test information function and conditional standard error of measurement are crucial to understanding the cause of systematic biases. As the *Standards* describe:



“The **standard error of measurement, as such, provides an indication of the expected level of random error over score points and replications for a specific population**. In many cases, it is useful to have estimates of the standard errors for individual examinees (or for examinees with scores in certain score ranges). These conditional standard errors are difficult to estimate directly, but can be estimated indirectly. For example, the test information functions based on IRT models can be used to estimate standard errors for different values of a latent ability parameter and/or for different observed scores” (p. 37).

The *Standards* also describe the purpose of calculating the test information function:



“The test information function, an important result of IRT, summarizes how well the test discriminates among individuals at various levels of ability on the trait being assessed... The information function may be viewed as **a mathematical statement of the precision of measurement at each level of the given trait**. The IRT information function is based on the results obtained on a specific occasion or in a specific context, and therefore it does not provide an indication of generalizability over occasions or contexts” (p. 38).

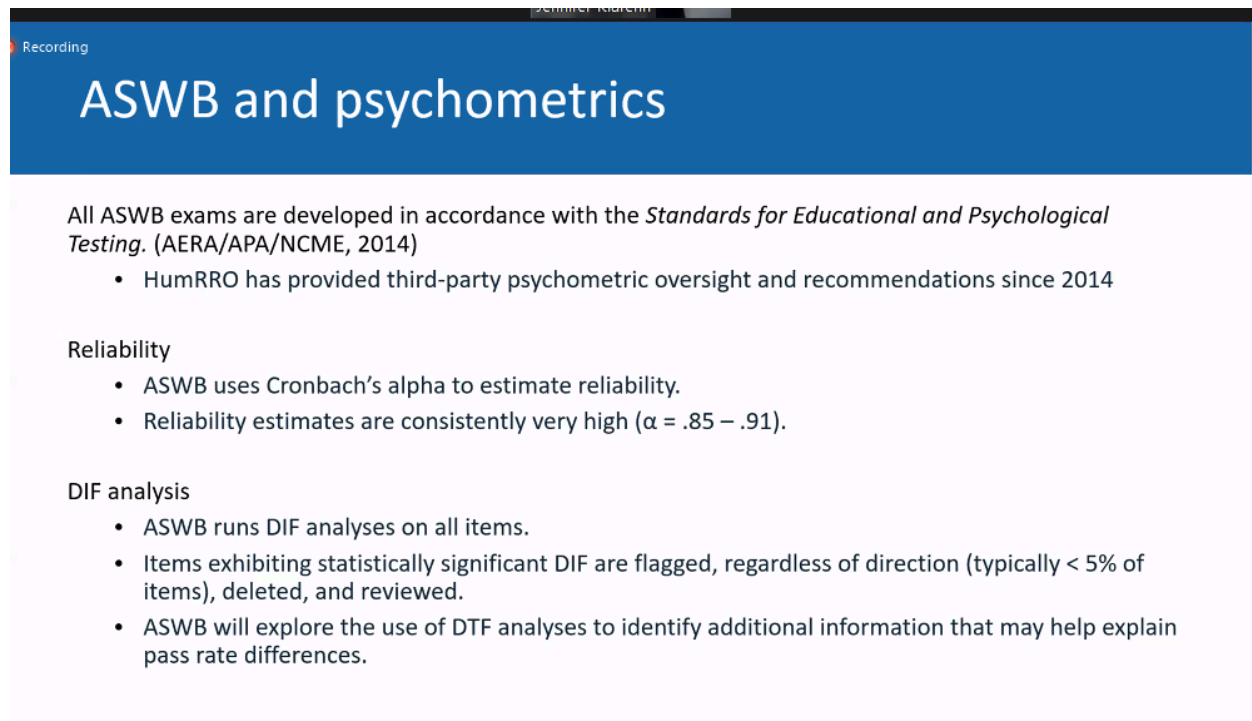
In 2023, ASWB rejected a (\$0) proposal from two social work psychometric experts and this author to perform these tests in the open source R stats program. ASWB did not provide access to their data to conduct the required psychometric validation tests.

Instead, ASWB assumes, without testing, that the conditional standard error of measurement and test information functions are equivalent across demographic groups. As the standards make clear, ASWB must provide evidence to support empirical assumption that these psychometric properties are equivalent across impacted groups.

ASWB Exams Are Inadequately Documented

ASWB provides inadequate documentation of their examination psychometrics, but even these small details reveal a clear explanation for how inequitable exams got worse.

In 2023, ASWB reported the only public psychometrics data since 2011. Here is a screenshot of the only slide in [a 60-minute presentation](#) that addressed validity & reliability:

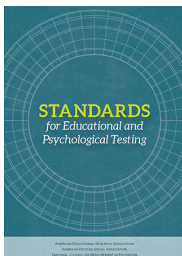


The screenshot shows a presentation slide with a blue header containing the title "ASWB and psychometrics". Below the header, the text reads: "All ASWB exams are developed in accordance with the *Standards for Educational and Psychological Testing*. (AERA/APA/NCME, 2014)". A bulleted list follows: "HumRRO has provided third-party psychometric oversight and recommendations since 2014". Under the heading "Reliability", there are two bullet points: "ASWB uses Cronbach's alpha to estimate reliability." and "Reliability estimates are consistently very high ($\alpha = .85 - .91$)". Under the heading "DIF analysis", there are three bullet points: "ASWB runs DIF analyses on all items.", "Items exhibiting statistically significant DIF are flagged, regardless of direction (typically < 5% of items), deleted, and reviewed.", and "ASWB will explore the use of DTF analyses to identify additional information that may help explain pass rate differences."

State boards do not receive any more information. According to a live ASWB's contract:

"ASWB will provide... the number of exams administered; the total number of items on the examination; the range of scores (lowest to highest); the mean and standard deviation taken from the annual technical report; and the Kuder-Richardson Formula 20 (K-R-20) Reliability Coefficient" (California BBS, p. 6).

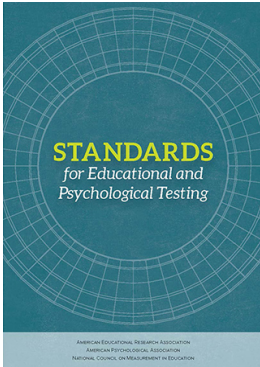
This is plainly inadequate documentation according to the *Standards* for documenting reliability.



The reporting of indices of reliability/precision alone—with little detail regarding the methods used to estimate the indices reported, the nature of the group from which the data were derived, and the conditions under which the data were obtained—**constitutes inadequate documentation**. General statements to the effect that a test is “reliable” or that it is “sufficiently reliable to permit interpretations of individual scores” are rarely, if ever, acceptable...test constructors and publishers are obligated to provide sufficient data to make informed judgments possible.

ASWB Clearly Violates Documentation Standards

ASWB is obligated to provide more psychometric data, as Standard 2.19 documents.

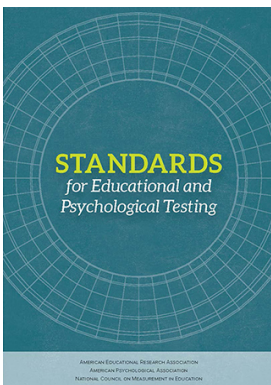


Standard 2.19: Each method of quantifying the reliability/precision of scores should be described clearly and expressed in terms of statistics appropriate to the method. The sampling procedures used to select test takers for reliability/precision analyses and the descriptive statistics on these samples, subject to privacy obligations where applicable, should be reported.

Comment: Information on the method of data collection, sample sizes, means, standard deviations, and demographic characteristics of the groups tested helps users judge the extent to which reported data apply to their own examinee populations...

Because there are many ways of estimating reliability/precision, and each is influenced by different sources of measurement error, **it is unacceptable to say simply, "The reliability/precision of scores on test X is .90."** A better statement would be, "The reliability coefficient of .90 reported for scores on test X was obtained by correlating scores from forms A and B, administered on successive days. The data were based on a sample of 400 10th-grade students from five middle-class suburban schools in New York State. The demographic breakdown of this group was as follows: . . .' (p. 47)

ASWB clearly does not follow the standards for documenting the reliability of its exams. This is a problem for reliably measuring social work competence across racial, cultural, linguistic, and age groups. Because it produces a high-stakes exam, documentation requirements are higher!



Standard 2.0: Appropriate evidence of reliability/precision should be provided for the interpretation for each intended score use.

Comment: The form of the evidence (reliability or generalizability coefficient, information function, conditional standard error, index of decision consistency) for reliability/precision should be appropriate for the intended uses of the scores, the population involved, and the psychometric models used to derive the scores. **A higher degree of reliability/precision is required for score uses that have more significant consequences for test takers.** Conversely, a lower degree may be acceptable where a decision based on the test score is reversible or dependent on corroboration from other sources of information.

Social Work Boards must enforce these standards.

ASWB Ruined Its Exams' Reliability

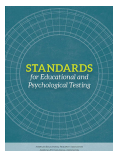
ASWB did not document why it stopped using Item Response Theory to measure exam reliability and started using Cronbach's Alpha.

In 2011, ASWB's prior psychometricians criticized the use of Cronbach's Alpha and attested that ASWB exams produced during that time were subject to decision consistency analysis:

" The ASWB examinations have shown high reliability estimates, in the nineties, both by the preferred advanced IRT model (decision consistency in pass/fail decisions) and the less relevant classical standards (KR-20, test reliability measure as shown by its internal consistency)" (p. 89)

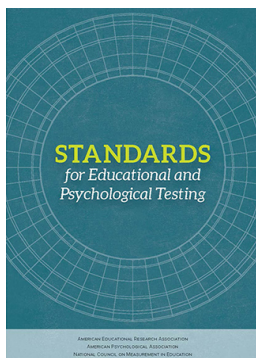
In 2024, ASWB uses less advanced reliability tests than in the 1990s.

ASWB's new psychometric consultants apparently do not agree with their psychometricians from the 1990s through the early 2010s that classical test theory measures are "less relevant" to ASWB exam reliability than decision consistency. Because ASWB does not adequately document its exams, we do not know why this change was made or its impact on the equivalence of exam scores. However, we do know they stopped using decision consistency, which violates the *Standards*.



Standard 2.16: **When a test or combination of measures is used to make classification decisions**, estimates should be provided of the **percentage of test takers who would be classified in the same way** on two replications of the procedure. (p. 46)

Instead of reporting decision consistency, *which ASWB did in the 1990s*, ASWB uses only internal consistency reliability. Moreover, ASWB is required by the *Standards* to report reliability for each relevant subgroup, such as those with different linguistic or cultural backgrounds:



Standard 2.11: Test publishers should provide estimates of reliability/precision as soon as feasible **for each relevant subgroup** for which the test is recommended.

Comment: Reporting estimates of reliability/precision for relevant subgroups is useful in many contexts, but it is especially important if the interpretation of scores involves within-group inferences (e.g., in terms of subgroup norms). For example, **test users who work with a specific linguistic and cultural subgroup** or with individuals who have a particular disability would **benefit from an estimate of the standard error for the subgroup**. (p. 45)

Broken Exams, Broken Workforce

In 2018, after new exams were introduced, the pass rate for *all social workers* reduced by about 10%. According to [ASWB’s Exam Pass Rate Analysis](#):

Figure 15. 2011–2021 Bachelors exam first-time pass rates by year and eventual pass rate

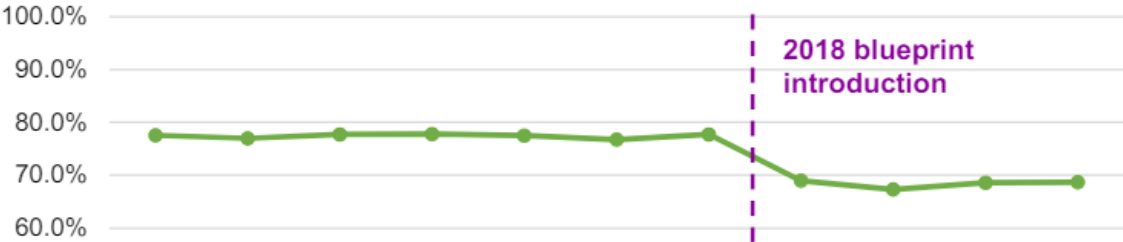
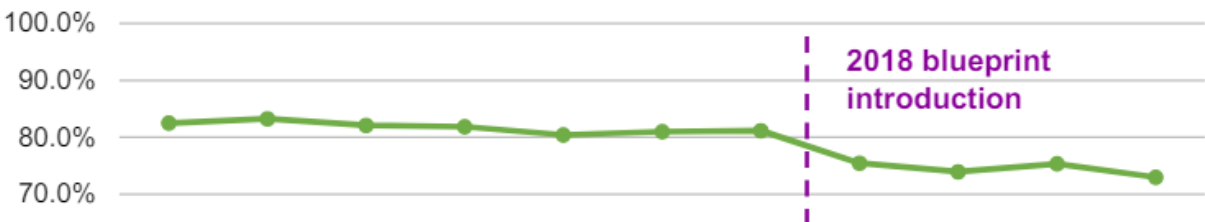


Figure 8. 2011–2021 Masters exam first-time pass rates by year and eventual pass rate



White Test-Takers Improve, Others Worsen

For the MSW exam, first-time and eventual pass rates demonstrate wide disparities across demographic groups from 2018-2021.

- White: 85-88% pass on their first time, regardless of age. Eventually, 91% pass.
- Black: 45% pass on their first time. 30% of over 50 pass first time. Eventually, 52% pass.
- Latine: 64% pass on their first time. 45% of over 50 pass first time. Eventually, 71% pass.
- English-secondary: 52% pass on their first time. Eventually 63% pass.

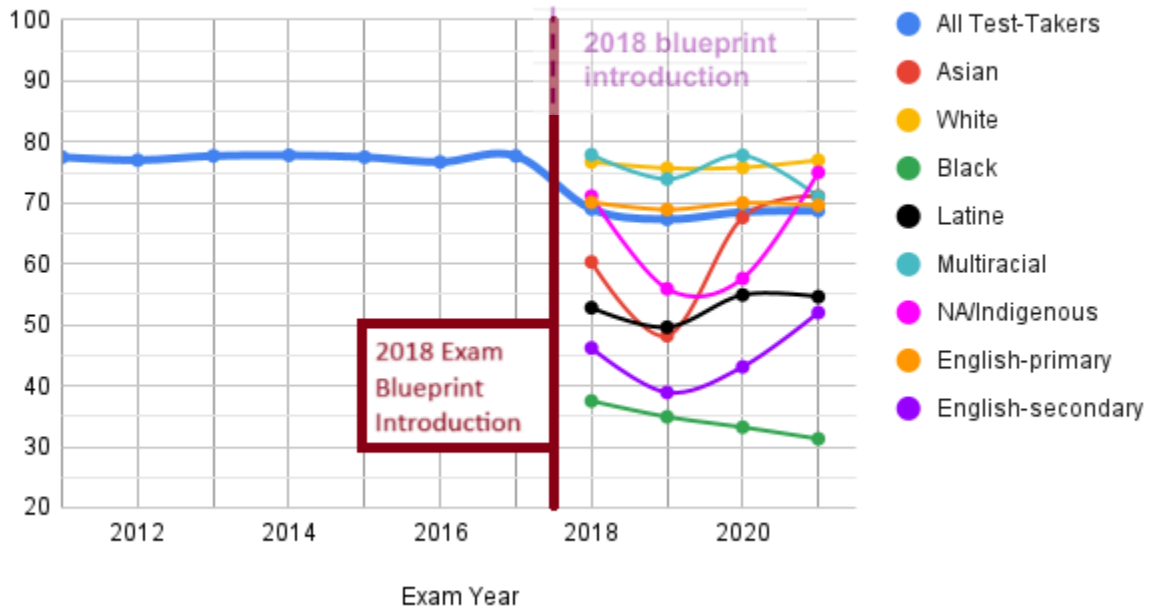
Because ASWB does not report equity data prior to 2018, it is impossible to know precisely how test-takers were impacted by the 2018 blueprint change.

But it is obvious these changes made exam inequity worse! After 2018:

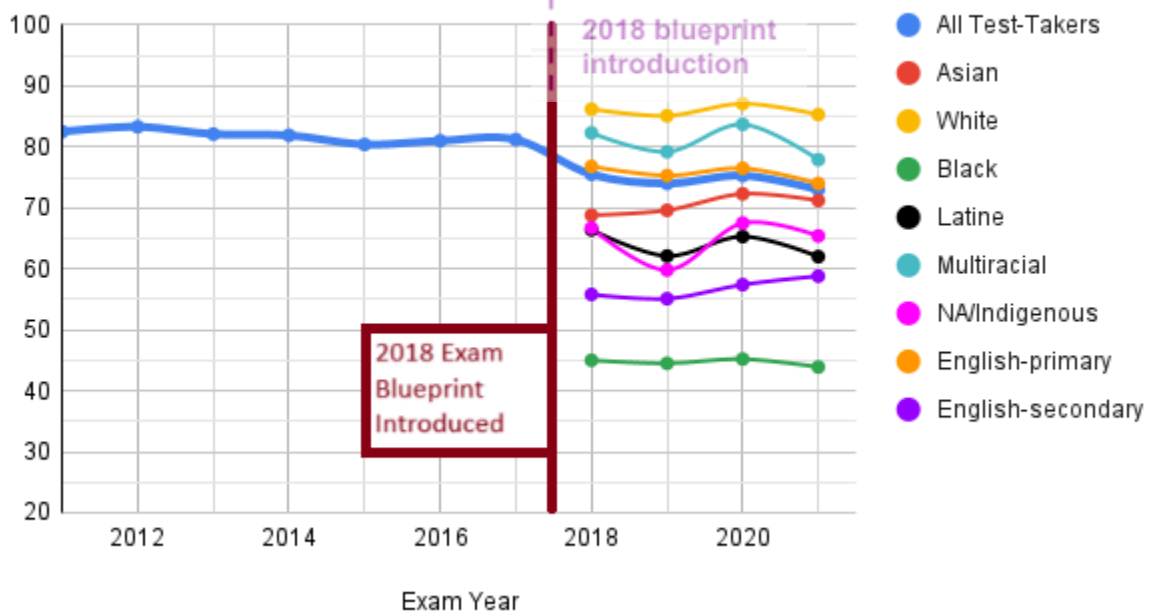
- White social workers’ test scores were at least 10% above the national average.
- Minoritized social workers’ scores were 20-40% less than the national average.

Here is a visual representation of the gap in exam scores reported by ASWB in their 2022 Exam Pass Rate Analysis by demographics. These indicate that white test-takers were least impacted by the 2018 exam blueprint and minoritized social workers were most impacted.

Impact of 2018 Blueprint on Bachelors Exam Pass Rates by Demographic



Impact of 2018 Blueprint on Masters Exam Pass Rates by Demographic



Likely, the 2018 examination blueprint made an already-inequitable exam less valid and fair.

Exam Pass Rates = Licensure Rates

Because exams are required for licensure, the workforce is missing the educated social workers who cannot pass the exams. Licensure and exam pass rates by demographic are highly similar.

Demographic group	ASWB Report First-time Pass Rate for MSW Exam	Senreich & Dale MSW Licensees in New York
White social workers (any age)	86%	78%
Black social workers (any age)	45%	48%
Latine social workers (any age)	64%	60%
Older Black social workers	30%	31%
Older Latine social workers	45%	39%
Older MSWs (any race)	65%	50%

When [Illinois](#) Eliminated the Exam

