



PSI EBRIEF #3

ITEM EXPOSURE COUNTERMEASURES USING CLASSICAL TEST THEORY

www.psonline.com



November 2014

Disclaimer. These guidelines are generalizations that are supported with evidence and do not represent an exhaustive list of supporting literature. Credentialing organizations should consult with a measurement professional since unique situations may require other acceptable evidenced based approaches and/or industry best practices. PSI believes this is a working document and any feedback is highly encouraged. Please send your feedback to marketresearch@psonline.com.

Factors to Consider	Summary	Supporting Literature
Use Multiple Forms	<p>Use more than one form of the exam.</p> <p>This addresses test security, which in turn, addresses validity.</p>	<p>Cohen, A. S., & Wollack, J. A. (2006). Test administration, security, scoring, and reporting. In Brennan, R. L. (Ed.), <i>Educational measurement (4th ed.)</i>. Westport, CT: Praeger. – Pages 363-364; 366-367</p>
Minimize Overlap Between Forms	<p>This is usually the case if you plan to post-equate, then you will need some intentional overlap, and the equating block will need to meet very specific parameters (see Kolen and Brennan, 2004).</p>	<p>Cohen, A. S., & Wollack, J. A. (2006). Test administration, security, scoring, and reporting. In Brennan, R. L. (Ed.), <i>Educational measurement (4th ed.)</i>. Westport, CT: Praeger. – Pages 363-364; 366-367</p> <p>Kolen, M. J., & Brennan, R. L. (2004). <i>Test equating, scaling, and linking: Methods and practices (2nd ed.)</i>. New York: Springer-Verlag.</p>
Testing Windows	<p>Use narrow testing windows</p> <p>Using narrow testing windows serves several purposes. First, it provides a level of security that continuous testing cannot provide. Second, it allows time to run non-standard analyses, such as data forensics.</p>	<p>Cohen, A. S., & Wollack, J. A. (2006). Test administration, security, scoring, and reporting. In Brennan, R. L. (Ed.), <i>Educational measurement (4th ed.)</i>. Westport, CT: Praeger. – Pages 363-364; 366-367</p>

<p>Replace/Dock Items</p>	<p>1) Monitor exposure rates and determine if item is no longer discriminating and many candidates have seen the item, it may need to be retired and/or replaced</p>	<p>1 / 2) Cohen, A. S., & Wollack, J. A. (2006). Test administration, security, scoring, and reporting. In Brennan, R. L. (Ed.), <i>Educational measurement (4th ed.)</i>. Westport, CT: Praeger.- Pages 363-364; 366-367</p>
	<p>2) "Replacing a large percentage [of items] between testing windows" p. 366</p>	
	<p>3) Retire items "Any times exposed to more than 15,000 examines must be retired (Way et al., 2002)</p>	<p>3) Way, W. D., Steffen, M., & Anderson, G. S. (2002). Developing, maintaining, and renewing inventory to support CBT. In C.N. Mills, M.T. Potenza, J.J., Fremer, & W.C. Ward (Eds.), <i>Computer-based testing: Building the foundation for future assessments</i>. Mahwah, NJ: Lawrence Erlbaum Associates. - Page 78</p>
	<p>On the other hand, docking items is another option where you temporarily restrict the use of the item for a certain amount of time.</p>	
<p>Restrict Most Frequently Selected Items When Updating/Building New Forms</p>	<p>See "Replace/Dock Items" above to determine if items should be replaced or docked that are used very frequently.</p>	<p>Davey, T. & Pitoniak, M. J. (2006). Designing computerized adaptive tests. In Downing, S. M. & Haladyna, T. M. (Eds), <i>Handbook of test development</i>. Mahwah, NJ: Lawrence Erlbaum Associates. - Page 553</p>

Factors to Consider	Summary	Supporting Literature
Use an exposure control parameter (more common on IRT/CAT exams)	<p>Value is between 0 and 1, and determines when an item would be used. For instance, an item that has an exposure control parameter of .10 is only selected 10% of the time to be on that exam form.</p> <p>A CTT guideline (or PIA under CAT) is to cap exposure at 0.2 to 0.25. Items that have exposure >0.25 should be removed from the bank before they become over-exposed.</p> <p>Any item over 0.4 should be retired from the bank.</p>	<p>1) Sympton, J. B., & Hetter, R. D. (1985). Controlling item-exposure rates in computerized adaptive testing. <i>Proceedings of the 27th Annual Meeting of the Military Testing Association</i> (pp. 973-977). San Diego, CA: Navy Personnel Research and Development Center.</p> <p>2) Georgiadou, E., Triantafillou, E., & Economides, A. A. (2007). A review of item exposure control strategies for computerized adaptive testing developed from 1983 to 2005. <i>The Journal of Technology, Learning, and Assessment</i>, 5, 5-37.</p> <p>3) Davey, T. & Pitoniak, M. J. (2006). Designing computerized adaptive tests. In Downing, S. M. & Haladyna, T. M. (Eds.), <i>Handbook of Test Development</i>. Mahwah, NJ: Lawrence Erlbaum Associates .- Page 554</p> <p>4) http://www.caveon.com/articles/df_article1)htm</p>

ABOUT PSI

PSI has over 70 years of experience providing worldwide testing solutions to corporations, federal and state government agencies, professional associations, certifying bodies and leading academic institutions. PSI offers a comprehensive solutions approach from test development to delivery to results processing, including pre-hire employment selection, managerial assessments, licensing and certification tests, distance learning testing, license management services and professional services. More information is available at www.psonline.com.

800.367.1565