From the beginning of the work of the Gordon Commission, staff members and Fellows have worked to compile a comprehensive collection and directory of the bibliographic resources used in the course of this work. Our Resource File is not a definitive collection; however, it does represent what we think of as the most important literature that has relevance for the work of the Gordon Commission. The collected works are organized under the working categories used by staff and can be searched using the "find" command found in the edit drop down menu (top bar in Adobe Acrobat), simultaneously using the "command" (⌘) and "f" keys on Macintosh computers, or simultaneously using the "ctrl" and "f" keys on Windows computers.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>History and Philosophy of Testing</td>
<td>3</td>
</tr>
<tr>
<td>Postmodern Test Theory</td>
<td>4</td>
</tr>
<tr>
<td>Validity</td>
<td>5</td>
</tr>
<tr>
<td>Validation/Validity Theory</td>
<td>5</td>
</tr>
<tr>
<td>Fairness</td>
<td>8</td>
</tr>
<tr>
<td>Evidence Centered Design</td>
<td>10</td>
</tr>
<tr>
<td>Consequences</td>
<td>12</td>
</tr>
<tr>
<td>Reliability and Measurement Error</td>
<td>13</td>
</tr>
<tr>
<td>Generalizability Theory</td>
<td>13</td>
</tr>
<tr>
<td>Classroom Assessment</td>
<td>13</td>
</tr>
<tr>
<td>Formative assessment</td>
<td>19</td>
</tr>
<tr>
<td>Performance Assessment</td>
<td>20</td>
</tr>
<tr>
<td>Stealth and Curriculum Embedded Assessment</td>
<td>23</td>
</tr>
<tr>
<td>Qualitative Analysis of Teaching and Learning Processes</td>
<td>26</td>
</tr>
<tr>
<td>Comprehensive Education</td>
<td>30</td>
</tr>
<tr>
<td>Large-Scale &amp; Summative Assessment</td>
<td>34</td>
</tr>
<tr>
<td>International Assessment</td>
<td>36</td>
</tr>
<tr>
<td>Cross-cultural/lingual Assessment (including Translation/Adaptation)</td>
<td>39</td>
</tr>
<tr>
<td>Relational Data Analysis and Management</td>
<td>42</td>
</tr>
<tr>
<td>To inform teaching</td>
<td>43</td>
</tr>
<tr>
<td>To address accountability</td>
<td>43</td>
</tr>
<tr>
<td>To guide learning</td>
<td>43</td>
</tr>
<tr>
<td>To direct orchestration</td>
<td>44</td>
</tr>
<tr>
<td>Reporting Assessment Results</td>
<td>44</td>
</tr>
<tr>
<td>Cognition and Assessment</td>
<td>44</td>
</tr>
<tr>
<td>Assessments and Technology</td>
<td>54</td>
</tr>
<tr>
<td>Individualization</td>
<td>60</td>
</tr>
<tr>
<td>Computer Adaptive Testing</td>
<td>61</td>
</tr>
<tr>
<td>Assessment of English Language Learners</td>
<td>63</td>
</tr>
<tr>
<td>Accommodations and Assessment of Students with Disabilities</td>
<td>66</td>
</tr>
<tr>
<td>Psychometrics and Psychometric Theory &amp; Principles</td>
<td>77</td>
</tr>
<tr>
<td>Item response theory</td>
<td>79</td>
</tr>
<tr>
<td>Differential item functioning</td>
<td>80</td>
</tr>
<tr>
<td>Evaluation and Policy Implementation</td>
<td>82</td>
</tr>
<tr>
<td>Accountability</td>
<td>91</td>
</tr>
<tr>
<td>Value-Added</td>
<td>96</td>
</tr>
<tr>
<td>Gaps between Research and Policy</td>
<td>97</td>
</tr>
<tr>
<td>Assessment and Context</td>
<td>97</td>
</tr>
<tr>
<td>Impact of Home-, School-, and District-level Variables on Student Achievement</td>
<td>97</td>
</tr>
<tr>
<td>Adaptive Human Behavior</td>
<td>100</td>
</tr>
<tr>
<td>Contextualism</td>
<td>101</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Situated Learning</td>
<td>101</td>
</tr>
<tr>
<td>Effervescence, Identity, Attribution, and Stereotype Threat</td>
<td>102</td>
</tr>
<tr>
<td>Measuring Capacity and Disposition (Content and Agency)</td>
<td>111</td>
</tr>
<tr>
<td>Interaction of Affective, Cognitive and Situative Processes</td>
<td>117</td>
</tr>
<tr>
<td>Social Emotional Learning</td>
<td>121</td>
</tr>
<tr>
<td>Other</td>
<td>127</td>
</tr>
</tbody>
</table>
History and Philosophy of Testing


**Postmodern Test Theory**


Validity

Validation/Validity Theory


**Fairness**


Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher


Sternberg, R. The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*. Volume 34, Issue 4, July-August 2006, Pages 321-350


**Evidence Centered Design**


**Consequences**


classrooms through social, academic and intellectual engagement First National Report. Toronto: Canadian Education Association.

Reliability and Measurement Error


Generalizability Theory


Classroom Assessment


Russ, J., Ehren, M. C. M., & Lesaux, N. (forthcoming). *Strategies teachers use to coach students to do well on the ELA state test: The case of New York City and Boston public elementary schools*.


**Formative assessment**


**Performance Assessment**


**Stealth and Curriculum Embedded Assessment**


**Qualitative Analysis of Teaching and Learning Processes**


Comprehensive Education


Large-Scale & Summative Assessment


International Assessment


First International Mathematics Study. (1964), from http://www.iea.nl/fims.html


Hill, C. (2001). Short-answer questions in testing reading comprehension in College English Research on Teaching College English in China (pp. 172–184). Beijing, China: Beijing
University Press.


Cross-cultural/lingual Assessment (including Translation/Adaptation).


College, Columbia University, New York.


Relational Data Analysis and Management


**To inform teaching.**


**To address accountability.**

information and assessment system for making evidence-based change at school, district, and policy levels. *Psychology in the Schools, 45*(3), 235-256.

**To guide learning.**


**To direct orchestration.**


**Reporting Assessment Results**


**Cognition and Assessment**


Aronowitz, R. (1984). Reading tests as texts. In D. Tannen (Ed.), *Coherence in spoken and*
written discourse (pp. 43-62). Norwood, NJ: Ablex


*Committee on Developments in the Science of Learning and Committee on Learning Research and Educational Practice, Commission on Behavioral and Social Sciences and Education.* Washington, DC: National Academies Press.


**Assessments and Technology**


Bennett, R. E. (2001). How the Internet Will Help Large-Scale Assessment Reinvent Itself. [Computer Assisted Testing; Educational Change; Internet; State Programs; Technological Advancement; Test Construction; Test Use; Testing Programs]. Education Policy Analysis Archives, 9(5), 1-23.


delivery system. *Teaching Exceptional Children, 42*(2), 6-12.


**Individualization**


Levin, I., Kojukhov, A. “Personalizing Education in Post-Industrial Society”, The Third International Conference on Digital Society, ICDS '09, Cancun, Mexico, pp. 20-23.


**Computer Adaptive Testing**


Assessment of English Language Learners


Accommodations and Assessment of Students with Disabilities


Center on Education Policy. (2009, November). Has progress been made in raising achievement for students with disabilities? Author: Washington, DC.


state policies on assessment participation and accommodations for students with disabilities *Synthesis Report 56*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.


Laitusis, C. C. (2010). Examining the Impact of Audio Presentation on Tests of Reading


Psychometrics and Psychometric Theory & Principles


**Item response theory**


**Differential item functioning**


Accommodated Versus Nonaccommodated Students. *Educational Assessment, 14*(1), 38-56. doi: 10.1080/10627190902816264


**Evaluation and Policy Implementation**


Harper Collins.


York: Springer-Verlag.


Public Law 107-110 - An act to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind, Pub. L. No. H.R. 1 C.F.R. (2002).


Accountability


Jacob, B. A. (2005). {Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools}. *Journal of Public Economics,* 89(5-6), 761-796. doi: citeulike-article-id:9172458


**Value-Added**


**Gaps between Research and Policy**


**Assessment and Context**

**Impact of Home-, School-, and District-level Variables on Student Achievement**


Stipek, D. Context Matters. *The Elementary School Journal, 0*(0), 000.


**Adaptive Human Behavior**


Murphy, R. (2010). Front Matter *Dynamic Assessment, Intelligence and Measurement* (pp. i-xvi): John Wiley & Sons, Ltd.
Contextualism


Situated Learning


Santa Monica, California.


**Effervescence, Identity, Attribution, and Stereotype Threat**


assessments for the 21st century (pp. 13–40). New York: Springer.


**Measuring Capacity and Disposition (Content and Agency)**


Leadership, 444–450.


**Interaction of Affective, Cognitive and Situative Processes**


consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, 64, 60–70.


Social Emotional Learning


doi: 10.1037/0033-295X.95.2.256


123


Other


Keppel, F. (1963, October 7). [Francis Keppel to Rice Clemow].

Keppel, F. (1963, August 10). [Francis Keppel to Joseph Califano].


Pearl, J. (2012). *In quest for the grammars of science*. Paper presented at the Celebration in Honor of Professor Judea Pearl, Winner of the 2011 Turing Award, UCLA Computer Science Department, Los Angeles.


Popovic, Z. (2012, August 20). DARPA Telecom


Political Writings (pp. 139-752). Indianapolis: Hackett Publishing Co.


Why Waldorf works. Everything you need to know about Waldorf education, from www.whywaldorfworks.org/


