

A New Tool to Strengthen Relational Teaching with Boys

Michael C. Reichert, PhD
michreich@comcast.net

Developing a Situational Judgment Test

Two global studies of successful practices in teaching boys were conducted between 2008-2012 by Rick Hawley and myself in partnership with the International Boys' Schools Coalition. Narratives were collected from over 2,000 teachers and 2,500 adolescent boys across approximately 40 schools in six countries, identifying successful teaching strategies with boys. Both groups agreed on the central feature in lessons deemed successful: "Relationship is the very *medium* through which successful teaching and learning is performed with boys," we reported.

As we presented the results of these studies at schools and conferences around the world, a common question arose: 'How can we enhance the relational skills of our present faculty?' Thinking about this question, I realized that a measure capable of providing reliable and valid feedback to teachers and coaches about their relational skills might be helpful. In addition to offering specific feedback to teachers, such a test would reinforce the necessity for a relational approach to boys for everyone.

To measure relational teaching, a test must recognize the subjective and situational nature of teaching and still do justice to boys' hard-wired human needs. To get at both domains, I was drawn to the situational judgment test, which has been used for over fifty years in personnel selection and professional development. These tests are built around realistic, hypothetical scenarios drawn from job analyses and assess how an individual will behave in certain common situations. Each situational judgment test is a bespoke tool, tailor-made to suit the role requirements of specific jobs. As part of the development process, subject matter experts (highly rated teachers, in this case) are asked to suggest effective and less effective solutions to common problems; a different group of subject matter experts then rate these responses from best to worst. Scores on situational judgment tests are most meaningful when they demonstrate a strong and reliable association with other measures of job success.

In partnership with Scotch College in Melbourne, Australia and Crescent School in Toronto, Canada, I formed a team in 2015 to develop a situational judgment test for educators of boys. My team included two experts in psychometrics and statistics associated with the University of Pennsylvania. We deduced a set of domains at the heart of successful relational teaching identified in our studies and drafted scenarios based on stories presenting common dilemmas in these domains. In cognitive focus group interviews with master teachers at two schools, Selwyn House School in Montreal and The Haverford School in Pennsylvania, these scenarios were reviewed for authenticity. In follow up interviews with another set of teachers from these schools, response options were created for each scenario.

The resultant items were then pilot tested with 100 teachers at 3 schools, Royal Saint George’s College in Toronto, The Haverford School and Selwyn House School, to gauge the overall reliability of the instrument. On the basis of item analyses from this pilot, the test was refined and further validation studies were conducted with teachers of Scotch College, Crescent School, Gilman Academy in Baltimore, Maryland, and Eaglebrook School in Massachusetts. Included in this round of testing were three other measures of effective teaching: a student evaluation, a value-added measure based upon government subject tests, and a self-evaluation.

In the final round of testing, a combined 26-item scale using standardized items from the self-evaluation and the 9-item situational test short produced an internal consistency reliability of 0.759. The test, in other words, could be trusted to produce a reliable and valid measure of teachers’ relational acumen.

Norming the Test

Next, we needed to create a set of norms against which test users could interpret their scores. Between February and May of this year, 21 IBSC member schools participated in a norming study. The sample included one school from South Asia, two schools from Africa, five schools from Australasia, three schools from Europe, and 10 schools from North America. Overall, 1,268 educators completed the test.

The test includes situational judgment (SJ) items and self-assessment (SA) items. The lowest possible score on the test is 13 and the highest score is 78. In this study, the raw average was 57.96 (*SD* = 6.38), with a minimum score of 27 and a maximum of 74, indicating the range of relational acumen found among teachers in this sample.

To make the test easy to interpret, the raw score of the SJTEB and each subscale was rescaled. Now a school (or an individual) can compare its performance against the norm of 50, knowing the standard deviation is 10.

Interpreting Test Results

Most importantly, educators and schools must find their results meaningful. To that end, results are reported in ways that let individuals and schools see quickly where they stand in relation to the test’s norms.

Below is a sample school report, comparing scores of teachers from School X, averaged across all 66 educators who took the test, to average scores of comparable groups. Results highlighted in yellow fall outside the norm of 50 to a significant degree. In a statistical sense, a difference of less than 2 points is considered “small,” a difference of 2-5 points is considered “medium,” and a difference of 5 or more points is considered “large.”

Sub-scale	Your	Norm Sample for Particular IBSC School
-----------	------	--

	School	Types		
		All	American	Secondary
Situational Judgment Questions	54.1	50	52.4	50.7
Self-report Questions	50.0	50	49.6	51.0
Combined Scale	51.5	50	50.5	51.1

For individuals, we created five levels of “relational teaching acumen,” from very low to very high, to allow individuals to see where they stand in relation to others and what their standing means.

Level	Scaled Score	% of norming sample respondents who were in this group	Typical performance
1 very low	<35	6.7%	Individuals in this group tended to select non-committal responses to survey items measuring relational attitudes (for example <i>neither agreeing nor disagreeing</i> with the statement “I never give up on a boy”). On judgment items, individuals in this group were unlikely to choose the most-relational option.
2 low	35-43	17.6%	Individuals in this group tended to select moderately relational responses to survey items (for example, often <i>agreeing</i> with the statement “I never give up on a boy”). On judgment items, individuals in this group had a good chance of choosing the most-relational option on the easiest questions.
3 medium	43-54	41.6%	In this, the largest group, individuals tended to select relational responses to survey items (for example, either <i>agreeing</i> or <i>strongly agreeing</i> with the statement “I never give up on a boy”). On judgment items, individuals in this group had a good chance of choosing the most-relational option on many of the question.
4 high	54-64.5	27.0%	Individuals in this group tended to select highly relational responses to almost all survey items (for example, <i>strongly agreeing</i> with the statement “I never

			give up on a boy”). On judgment items, individuals in this group had a good chance of choosing most-relational option on all but the single most difficult question.
5 very high	>=64. 5	7.2%	Individuals in this group tended to select highly relational responses to even the most challenging survey items (for example, most likely to <i>strongly disagree</i> with the statement “There are some boys, particularly those experiencing difficulties beyond my classroom, whose attitudes to learning and academic outcomes are beyond my control”). On judgment items, individuals in this group had a good chance of choosing the most-relational option on even the most difficult question.

We also contrast the number and percent of individuals from a particular school at each level with the percent of teachers in the IBSC norming sample at that level. What School X can see, for example, is that more of its teachers scored in the average range and fewer scored low and very low.

	Your School		IBSC Norming Sample
	N	%	%
Very Low	1	1.5%	6.7%
Low	8	12.1%	17.6%
Medium	36	54.5%	41.6%
High	17	25.8%	27.0%
Very High	4	6.1%	7.2%

Next Steps

The Situational Judgment Test for Educators of Boys is an authentic, statistically sound, and challenging tool that can be used to foster self-reflection and professional discussion about the primacy of teaching relationships. The test can help school faculty engage with each in lively and productive conversations as they consider their own choices – and the values they reflect - in relation to those of their colleagues and against the backdrop of a broad sample of teachers of boys. On questions where they score outside of the mean, teachers can hear their colleagues’ reasons for answering as they did and consider the reactions and beliefs underlying their own choices.

While we digest the various studies we have conducted and consider next steps, we want to make the test available to IBSC member schools in recognition of the central role the Coalition has played in its development. To that end, we make the following offer: any member school wishing to administer the test can arrange with us to do so. For IBSC members we will offer a discounted price, simply to cover administrative, reporting and additional research expenses, for six months from September 1, 2018.

Some schools seeking to take advantage of this offer may wish for more guidance about what test results mean beyond the score report. We offer two additional options. First we will create an online slide show, with voiceover, of about 20 minutes length to explain what is meant by relational acumen, how the concept arose from research conducted with boys and teachers, and what individuals as well as schools can take from their results. To help schools review their results in a more interactive way, we can also conduct an in-person workshop.

Schools interested in administering the test or in either of the follow-up options can contact me directly.

In closing, I am deeply appreciative of the partnership the IBSC has offered in this line of research from the beginning. Friendship, collegial conversation, and support for the effort to get things right for boys have been uniform features of my experience at member schools and conferences. In the development of this measure of relational judgment, as with all of my research, the IBSC has been a critical, valued partner.