

Technical Report: October 2021 CHRP-KE

HR | Human Resources
PA | Professionals Association

16 November 2021



Contents

Executive Summary	5
Administration.....	6
Form Setting	6
Testing Window	7
Analysis	8
Data Cleaning and Integrity Checks.....	8
Post-Examination Survey.....	10
Initial Analysis	13
Key Validation.....	15
Establishing the Pass Mark: Equating.....	16
Scoring	25
Key Examination Metrics.....	27
Related Development Activities	28
Item Writing	28
Appendix A.....	30
Appendix B.....	32

List of Tables

Table 1: Test forms as administered	7
Table 2: Content-related post-examination survey questions*	10
Table 3: Administration-related post-examination survey questions*	11
Table 4: Testing location	11
Table 5: Preferred location (live remove proctoring candidates)	12
Table 6: Reason for choosing own location (live remove proctoring candidates)	12
Table 7: Evaluation of testing experience (live remove proctoring candidates)	12
Table 8: Value in future candidates being able to test from their own location (live remote proctoring candidates)	12
Table 9: Able to write at a convenient location (test centre candidates)	13
Table 10: Value in future candidates being able to test from their own location (test centre candidates)	13
Table 11: Initial examination statistics – Combined across forms	14
Table 12: Section item statistics	14
Table 13: CHRP Examination Validation Committee members – Key validation	15
Table 14: Final scored examination fit to blueprint	16
Table 15: Anchor item fit to blueprint – To June 2021	18
Table 16: Equating parameter table – Total pass mark, to June 2021	18
Table 17: Equating outcome table – Total pass mark, to June 2021	19
Table 18: Anchor item fit to blueprint – To February 2021	20
Table 19: Equating parameter table – Total pass mark, to February 2021	20
Table 20: Equating outcome table – Total pass mark, to February 2021	21
Table 21: Anchor item fit to blueprint – To August 2020	22
Table 22: Equating parameter table – Total pass mark, to August 2020	22
Table 23: Equating outcome table – Total pass mark, to August 2020	23
Table 24: Equating outcome table – Combined results	23
Table 25: Historical pass rates	24
Table 26: CHRP Examination Validation Committee members – Pass mark approval	24
Table 27: Total and functional area scores for all candidates	25
Table 28: Correlations between functional area scores for all candidates	26
Table 29: Key examination metrics – Candidates included in analysis only	27

Table 30: Item writers.....28

Table 31: CHRP-KE Blueprint structural variables30

Table 32: Functional area weights on the CHRP-KE31

Table 33: Competencies not eligible on the CHRP-KE.....31

List of Figures

Figure 1: Examination time distribution for all candidates 9

Figure 2: Candidate volume and score trends across testing window 9

Figure 3: Score distribution for all candidates.....26

Executive Summary¹

Note that this technical report covers only the primary new form or forms administered during an administration, and not detailed results for all forms used (which may include previously used forms, scrambled forms, and other modifications to maintain exam and score integrity).

The CHRP-Knowledge Exam (CHRP-KE²) was administered to 394³ candidates using computer-based testing and live remote proctoring October 19–November 2, 2021, inclusive. The examination comprised 175 four-option multiple choice items and had a 3½-hour time limit.

As per the CHRP-KE blueprint, the exam was scored using the 145–155 best-performing items (while adhering to the prescribed distribution across functional areas). The mean score for first-time candidates⁴ ($n=295$) was 102.7 (68.0%), and for all candidates it was 99.9 (66.1%), out of 151 scored items. Reliability was strong at .90. The final set of scored items adhered to the blueprint parameters.

The pass mark was set using equating back to the June 2021, February 2021, and August 2020 administrations, yielding an integer pass mark of 95. Equating was conducted to compensate for minor changes in exam form difficulty so that any given candidate has an equivalent hurdle regardless of when they write the CHRP-KE. This pass mark resulted in a pass rate for first-time candidates of 70.5% and a pass rate for all candidates of 62.4%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards⁵ and ISO 17024 standards.⁶

¹ This technical report is an abbreviated version of the full report. Information has been excluded that if known to candidates could negatively affect the validity of future candidate test score interpretations. This includes item-level statistics, some information about the construction of test forms, and some specific details concerning equating.

² The CHRP-KE was titled the CKE 1 up until the Fall of 2020. Any reference in this report to past administrations of the CHRP-KE will use the new title.

³ In the initial analyses, data were available on 392 candidates (of which 294 were part of the first-time candidate group), and so some values in this report are based on that slightly smaller candidate sample. Equating was calculated using the full sample of 295 first-time candidates. One candidate, who was not included in the first-time candidate group, later had their results nullified due to an administration irregularity but appears in tables in this report as part of the full cohort of 394 candidates.

⁴ Excludes those who had failed an HRP examination in the past, who were identified as being statistical outliers, or who had written an alternative test form.

⁵ National Commission for Certifying Agencies (2014). *Standards for the accreditation of certification programs*. Washington, DC: Institute for Credentialing Excellence.

⁶ International Organization for Standardization (2012). *ISO/IEC 17024:2012 Conformity assessment – General requirements for bodies operating certification of persons*. Geneva: International Organization for Standardization.

Administration

Form Setting

Using only validated test items, Wickett Measurement Systems prepared three 175-item test forms (using a combination of scored and experimental test items). Wickett constructed the final test forms according to the following parameters:

1. Including only items validated by the validation panel in the past 3 years
2. Fitting the total item count of 175
3. Excluding enemy items
4. Matching the blueprint target value (+/- 2%) for each functional area
5. Maximizing spread across competencies
6. Reducing item exposure
7. Selecting items with perceived psychometric effectiveness, using statistics from previous administrations as available

Wickett proofed the final forms for text errors and detection of potential enemy items. Items flagged as enemies were replaced.

After selecting the 175 items for each form, Wickett split the forms in half to allow for the administration of the exam in two sections. Section 1 was allocated 88 items and Section 2 was allocated 87 items. With each form, the two sections were set to balance for:

- Number of words
- Time per item
- Item difficulty
- Item discrimination (adjusted point-biserial)
- Number of experimental items
- Adherence to blueprint
- Number of anchor items

The final form composition for the October CHRP-KE forms is shown in Table 1. All functional areas are within the limits of their targets, and therefore the forms reflect the blueprint (see Appendix A for the CHRP-KE blueprint). Differences between targets and actuals reflects differential allocation of experimental items rather than a deviation from scored item targets.

Note that at any administration, HRPAs also makes use of previously validated and administered test forms along with new test forms, in addition to employing other mechanisms to maintain the integrity of the exams and candidate scores.

Table 1: Test forms as administered

	Functional Area	Actual Items	Target
10	Strategy	7–8	6–8
20	Professional Practice	19–20	18–21
30	Organizational Effectiveness	23–24	21–24
40	Workforce Planning & Talent Management	22	21–24
50	Labour & Employee Relations	20–21	18–21
60	Total Rewards	21–22	21–24
70	Learning & Development	23	21–24
80	Health, Wellness & Safe Workplace	18–19	18–21
90	HR Metrics, Reporting & Financial Management	19–20	18–21
	TOTAL	175	175

Testing Window

The examination was administered via computer-based testing using live remote proctoring and at Prometric test sites primarily in Ontario. The testing window was October 19–November 2, 2021, inclusive, and 392 candidates wrote the exam⁷.

Candidates were able to select either a test centre (assuming one was available reasonably close to them) or live remote proctoring from a location of their choosing. As a result of continued closures, most candidates wrote using live remote proctoring. Standard security methods (as per Prometric protocols⁸) were employed for both methods. Candidates were allowed one 15-minute break after submitting section 1 and before beginning section 2. This break did not count against total time for the candidate.

Candidates had access to a basic-function calculator on screen. No other aids or resources were allowed.

⁷ Due to technical difficulties requiring the rescheduling of some candidates, testing continued through to November 3, 2021.

⁸ Information on procedures and security can be found at www.prometric.com/ProProctor and www.prometric.com/proproctorcandidate.

Analysis

Data Cleaning and Integrity Checks

Prometric provided data in .xml format via a secure ftp site. Candidate files were provided as candidates completed the examination throughout the testing window. These files were extracted to Microsoft Excel for processing. They contained identifying information for each candidate, form information, start and stop times, answer string, key string, candidate total score, item comments if the candidate made any, and time spent per item.

The data files received were reconciled against the roster provided by Prometric and HRP to ensure that all .xml files had been received. Further, each candidate total score as computed by Prometric was reconciled with that computed by Wickett for the full set of 175 items to verify key accuracy. Comments on items were also reviewed to identify any specific item-level issues. No problems were encountered.

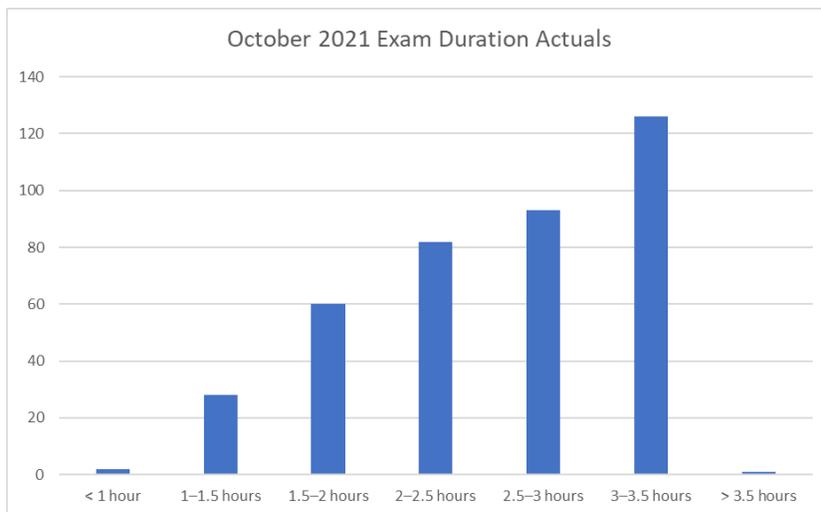
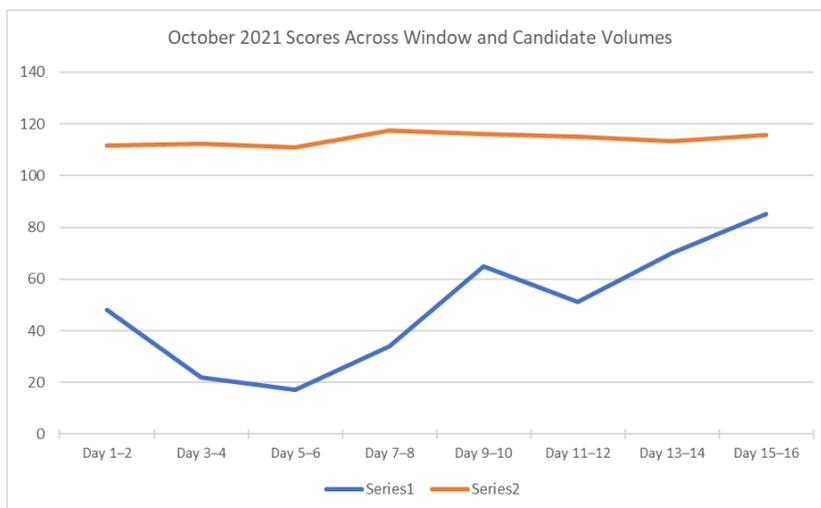
The average time taken by all candidates was assessed to detect potential examination timing concerns. The distribution is shown in Figure 1. The mean was 2 hours, 33 minutes (8 minutes more than in June 2021; on average, form A candidates took 2 hours, 35 minutes, form B candidates took 2 hours, 29 minutes, and form C candidates took 2 hours, 35 minutes). The time limit on the CHRP-KE was 3½ hours, suggesting that time was not a factor in scores across candidates. One candidate who was granted additional time as a testing accommodation exceeded the regular time limit of 3½ hours.

Twenty-four candidates (6%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 8 candidates (2.0%) left at least 1 item blank, suggesting that those candidates timed out of the exam before being able to complete it. These metrics will continue to be monitored, but at present do not appear problematically high.

The correlation between scores on the 175 items and time spent writing the examination was negligible at a value of .01 for form A, negligible at a value of .07 for form B, and small at a value of -.16 for form C, suggesting that time constraints did not generally have an impact on candidate performance.

Candidate scores across the window were computed to look for any evidence of item exposure. As shown in Figure 2, there was little variation across the window. The difference between scores for candidates writing in the first 2 days and those writing in the last 2 days was an increase of 4.3 marks out of 175.

As a matter of interest, candidate volumes were also examined across the window; these are also shown in Figure 2. Though not psychometrically meaningful, there is a pattern for candidates to prefer to book towards the end of the window rather than the start.

Figure 1: Examination time distribution for all candidates**Figure 2: Candidate volume and score trends across testing window**

After removing candidates who were administered a previously used test form (who were scored using the same decisions employed at the time that form was originally used), scores were calculated for all remaining candidates based on the full set of 175 items. No candidates were flagged for an abnormally low or high score (z value outside ± 3.0). Also, the 175 items were arbitrarily broken into 7 blocks of 25 items for each candidate; the 7 resulting subscores for each candidate were evaluated for outliers as well. For candidates with any subscore more than 3 standard deviations (SD) from their average z -score, the .xml file was examined closely for any issues. Candidates who left 5 or more blanks were also flagged for removal from analysis (no candidates were flagged on this criterion). As a result of all of these factors, 0 candidates were removed from analysis.

Candidates who had failed a previous HRP examination (CKE, CHRP-KE, or CHRL-KE) scored lower than did those who had not (59.8% and 67.3%, respectively, on the full exam of 175 items). This difference was meaningful and significant ($t(255)=7.95$, $p<.001$). In keeping

with standard procedures, these candidates were removed from subsequent analyses. The CHRP-KE analysis proceeded with 294 candidates.

Owing to the modest number of candidates, all subsequent analyses were interpreted with caution.

Post-Examination Survey

Candidates were provided with access to the post-examination survey immediately after submitting their responses to the CHRP-KE; 389 responses were obtained from candidates (response rate, 99.2%).

Table 2 shows the content-related questions; there was a tendency to neutrality on these questions. The rating for perceived fairness (Question 8) warrants monitoring as it continues to be low. Table 3 shows the responses to the administration-related questions. Note that candidates were generally very positive about the administration experience.

Table 2: Content-related post-examination survey questions*

	Question	SA	A	N	D	SD	Score	Agreement	Agreement last 5[^]
1.	The time allotted for this examination was sufficient.	227	125	16	16	5	4.5	90%	91%
2.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	124	159	61	41	6	4.1	72%	71%
3.	I feel I was adequately prepared to write this examination.	14	145	160	62	10	3.6	41%	50%
4.	The questions in the examination were clearly written.	47	204	97	40	3	3.9	64%	65%
5.	The terminology used in the examination was accurate.	44	234	87	20	4	4.0	71%	71%
6.	The situations presented in the examination were realistic.	64	242	69	9	2	4.1	79%	80%
7.	The questions in the examination reflected the examination blueprint.	35	168	139	43	4	3.8	52%	55%
8.	The examination was a fair assessment of my ability.	26	132	133	79	17	3.5	41%	47%

*Response categories: SA = strongly agree; A = agree; N = neutral; D = disagree; SD = strongly disagree.

[^]Mean value of candidate agreement across the previous 5 administrations.

Table 3: Administration-related post-examination survey questions*

	Question	SA	A	N	D	SD	Score	Agreement	Agreement last 5[^]
9.	I was able to book to write the examination at a time that was convenient for me.	152	163	34	31	8	4.2	81%	86%
10.	I was well informed about the examination rules and regulations.	225	146	9	6	3	4.5	95%	96%
11.	Proctors enforced the exam-day rules.	255	115	15	1	0	4.7	96%	97%
12.	Proctors were professional and courteous.	265	103	16	3	1	4.7	95%	96%
13.	The tutorial helped me understand how to complete the examination on the computer.	213	135	33	3	3	4.5	90%	93%
14.	Navigation through the examination was easy and intuitive.	210	159	15	2	1	4.5	95%	96%

*Response categories: SA = strongly agree; A = agree; N = neutral; D = disagree; SD = strongly disagree.

[^]Mean value of candidate agreement across the previous 5 administrations.

Candidates were asked where they wrote the examination and based on their response the questions that followed differed (see Table 4). Restrictions in Ontario at the time of the testing meant very few test centres were open. The one individual who responded as writing at a test centre likely did so in error.

Table 4: Testing location

Response	Count	%
Test centre	29	7%
Own location	358	93%

Candidates who indicated they tested in the own location (via live remote proctoring) responded to questions shown in Table 5 through Table 8. These candidates favoured using their own location versus a test centre, were generally positive about the experience and felt that HRP should continue to offer the option in the future. As expected, COVID-19 related concerns and closures were a motivating factor for many in choosing live remote proctoring, though a substantial portion identified the convenience of not having to travel as a main driver.

Table 5: Preferred location (live remove proctoring candidates)

Response	Count	%
I preferred using my own location.	238	67%
I preferred going to a test centre.	42	12%
I have no preference.	75	21%

Table 6: Reason for choosing own location (live remove proctoring candidates)

Response	Count	%
No test centres were open in my area.	103	29%
I preferred to avoid being around other people.	39	11%
I liked the convenience of not having to travel to a test centre.	152	43%
I felt like I would perform better in my own environment.	39	11%
Other (please specify)	24	7%

Table 7: Evaluation of testing experience (live remove proctoring candidates)

	Count	%
Very positive	123	35%
Positive	153	43%
Neutral	66	19%
Negative	8	2%
Very negative	5	1%

Table 8: Value in future candidates being able to test from their own location (live remote proctoring candidates)

Response	Count	%
Yes	352	99%
No	4	1%

Candidates who indicated they tested in a test centre responded as shown in Table 9 and Table 10. These candidates were positive about being able to write at a convenient location, and were also supportive of HRPA continuing to offer the option of writing using live remote proctoring in the future.

Table 9: Able to write at a convenient location (test centre candidates)

	Count	%
Strongly agree	10	34%
Agree	11	38%
Neither agree nor disagree	0	0%
Disagree	6	21%
Strongly disagree	2	7%

Table 10: Value in future candidates being able to test from their own location (test centre candidates)

Response	Count	%
Yes	24	83%
No	5	17%

Open-ended questions were also posed to candidates asking for any additional comments in general and regarding test delivery method. Those comments were provided to HRP A for information and consideration. Nothing actionable with respect to scoring emerged in these comments.

Initial Analysis

The full CHRP-KE examination was 175 items, of which approximately 150 were to be scored. The other 20–30 items were not intended to be scored. Across the 3 new forms, 151 items were available for scoring on each, after removing items designated as experimental.

The initial analysis summary statistics are presented in Table 11 (the previous administration values are also provided as a point of reference). The section statistics are presented in Table 12.

Table 11: Initial examination statistics – Combined across forms

Index	Oct. 2021	Jun. 2021	Feb. 2021
Items	151	151	151
Total candidates	392	342	316
Candidates in analysis	294	267	236
Mean score	102.7 (68.0%)	106.8 (70.7%)	106.5 (70.5%)
Standard deviation	16.7	17.1	16.7
Score range	56–137 (37.1–90.7%)	52–141 (34.4–93.4%)	56–138 (37.1–91.4%)
Cronbach's alpha	.90	.91	.90
Mean r_{pb}^*	.23	.24	.23

Table 12: Section item statistics

Index	Section 1	Section 2
Total items	88	87
Scored items	76	75
Candidates in analysis	294	
Mean	50.8 (66.9%)	51.9 (69.2%)
Standard deviation	8.8	8.7
Range	28–70	24–70

A simple comparison between scores obtained by test centre candidates (mean score of 63.6%) and live remote proctoring candidates (mean score of 66.4%) was made to evaluate if there was any problematic difference in performance. The small number of candidates means this analysis is inconclusive, but a lack of significant effect ($t(390)=1.34$, ns) is at least supportive of there being no overall difference in candidate performance. There is a leaning towards higher scores from live remote proctoring candidates that will be monitored across future administrations.

Though not reported here, several additional analyses were added with administration to investigate potential candidate misconduct. These results were reported confidentially to HRP.

Standard classical test theory analysis was conducted to identify the following:

1. Item difficulty (percent obtaining correct result, p)

2. Item discrimination (corrected point-biserials, r_{pb}^*)
3. Distractor quality (based primarily on distractor discrimination)

Wickett compiled these statistics, along with any comments made by candidates concerning specific items, to identify items that may have been keyed incorrectly or that were performing poorly. Most emphasis was placed on corrected point-biserials as evidence of item quality and on difficulty through removal of ineffective very easy or very hard items. Items were ranked from worst performing to best performing accordingly.

Key Validation

Key validation was conducted via web meeting on November 5, 2021, using members of the CHRP Examination Validation Committee (EVC). The EVC (Table 13) was reminded of basic item and test analysis methods and was oriented to the main statistics used to evaluate the quality of the CHRP-KE.

Table 13: CHRP Examination Validation Committee members – Key validation

Member	Credential	Years of Relevant Experience	Joined EVC	Industry
Sunday Ajao	CHRL	15–20	2017	Banking/Finance
✓ Nancy Brandon	CHRL	20–25	2021	Power and Utilities
Roxanne Chartrand	CHRL	20–29	2018	Insurance
✓ Claire Chester	CHRL	10–15	2017	Long term care facility
✓ Patrizia Finucan	CHRL	10–15	2021	Education
✓ Tanya Gopaul	CHRL	10–15	2017	Banking
Annette Lawrence	CHRL	5–10	2021	Non-profit
Suman Seth	CHRL	15–19	2018	Public sector/education
Kriss Stone	CHRL	10–15	2017	
✓ Michelle Sultan	CHRL	10–15	2021	Education
Ileean Tait	CHRL	15–20	2017	Environmental
Patricia Verkley	CHRL	10–15	2019	Not-for-profit
Karen Weiler	CHRL	20–29	2017	Software/ Communications

✓ Participated in the session.

The committee was informed that test reliability, as measured by Cronbach's alpha, was .90 based on the set of 151 potentially scored items and that this was well above the generally accepted threshold of .80.

The committee was informed that 3 items fell below the flagging criteria and 3 items were flagged for review based on candidate comments, and they reviewed each of these for content accuracy and clarity. The committee removed none of these items from scoring, and the set of 151 items was approved for use in scoring the October 2021 CHRP-KE candidates who took this form.

The group also reviewed and made decisions about the future use of experimental items in this session.

Not all remaining items were strong-performing, and several items were retained that were easy or hard or that had a low corrected point-biserial in this sample of candidates. Most were moderate to strong items, however. The final alpha for the set of 151 scored items was .90. The difficulties ranged from 28.8% to 93.6%, with a mean of 68.0%. The r_{pb}^* values ranged from $-.08$ to $.49$, with a mean of $.23$.

Table 14 presents the scored CHRP-KE's final fit to the examination blueprint. In all cases, the final number of scored items in a functional area fit within the established range.

Table 14: Final scored examination fit to blueprint

	Functional Area	Actual	Min.	Target*	Max.	Blueprint Range
10	Strategy	6	5	6	7	4% ± 1%
20	Professional Practice	17	14	17	19	11% ± 2%
30	Organizational Effectiveness	19	17	20	22	13% ± 2%
40	Workforce Planning & Talent Management	19	17	20	22	13% ± 2%
50	Labour & Employee Relations	17	14	17	19	11% ± 2%
60	Total Rewards	19	17	20	22	13% ± 2%
70	Learning & Development	20	17	20	22	13% ± 2%
80	Health, Wellness & Safe Workplace	17	14	17	19	11% ± 2%
90	HR Metrics, Reporting & Financial Management	17	14	17	19	11% ± 2%
	Total	151				

*Adds to 154 due to rounding.

Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014),⁹ was used to establish the pass mark for the October 2021 CHRP-KE. The goal of this process was to set a pass mark for the October 2021 CHRP-KE that would be equivalent to that set for previous CHRP-KE administrations; that is, to

⁹ Kolen, M.J., & Brennan, R.L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer.

set a pass mark that would give each candidate the same probability of passing regardless of which form they took.

The passing standard for the CHRP-KE was originally set after the November 2015 offering of the CHRP-KE using the Modified Angoff method. General details on that method can be found in Appendix B. Specific information on the standard setting session is provided in the technical report issued for the November 2015 administration.

Three equating procedures were conducted back to different administrations (June 2021, February 2021, and August 2020). Separate procedures were conducted to reduce the effects of sample variability and arrive at the most accurate equated pass mark.

Equating Back to the June 2021 Administration

Linear equating was the chosen method for setting the pass mark. Linear equating is preferred with more than 100 candidates, and equipercentile equating is preferred with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc¹⁰ equating is most prudent.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 SD units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form (and equivalent within each functional area) that adhered to the blueprint. Items with an increase or decrease of 10% in terms of difficulty were also removed as anchors. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.68 and a mean corrected point-biserial of .25 (for October 2021 candidates).

Table 21 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are well-aligned with targets and reflect the scope and approximate weighting across the full exam.

¹⁰ Kim, S., & Livingston, S.A. (2010). Comparisons among small sample equating methods in a common-item design. *Journal of Educational Measurement*, 47, 286-298.

Table 15: Anchor item fit to blueprint – To June 2021

Area*	Actual	Target
10	4%	4%
20	11%	11%
30	11%	13%
40	13%	13%
50	9%	11%
60	13%	13%
70	13%	13%
80	13%	11%
90	11%	11%

*See Table 14 for the full name of each functional area.

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker equating was considered the preferred method.

Table 22 shows some of the parameters used to derive the equating estimates, along with other parameters describing the test forms. Of note is that on the anchor items, the candidates taking the October 2021 CHRP-KE scored slightly lower than the candidates taking the June 2021 CHRP-KE (68.0% vs. 69.0%; $t(560)=0.85$, *ns*). Because the October 2021 CHRP-KE candidates scored lower (non-significance notwithstanding), they would likely have a slightly lower pass rate as compared to June 2021 candidates.

The equating analysis bears this out (Table 23). All methods indicate a pass mark of 94 or 95, with the preferred Tucker method providing a value of 94. The pass rate based on this equating run is lower, as expected, than what was seen in June 2021. The Tucker equating value of 93.86 was extracted from this analysis for use in setting the final pass mark.

Table 16: Equating parameter table – Total pass mark, to June 2021

		Jun. 2021	Oct. 2021
N		267	295
Scored items		150	151
Mean score	Total	70.7%	68.0%
	Anchors	69.0%	68.0%

Table 17: Equating outcome table – Total pass mark, to June 2021

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First-time
Equating Jun. 2021	96.02	97	65.5%	72.7%
Tucker	93.86	94	64.5%	71.9%
Levine observed	94.00	95	62.4%	70.5%
Mean	93.77	94	64.5%	71.9%
Circle Arc 1	93.66	94	64.5%	71.9%
Circle Arc 2	93.63	94	64.5%	71.9%

Equating Back to the February 2021 Administration

Linear equating was the chosen method for setting the pass mark. Linear equating is preferred with more than 100 candidates, and equipercentile equating is preferred with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc¹¹ equating is most prudent.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 SD units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form (and equivalent within each functional area) that adhered to the blueprint. Items with an increase or decrease of 10% in terms of difficulty were also removed as anchors. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.68 and a mean corrected point-biserial of .25 (for October 2021 candidates).

Table 21 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are well-aligned with targets and reflect the scope and approximate weighting across the full exam.

¹¹ Kim, S., & Livingston, S.A. (2010). Comparisons among small sample equating methods in a common-item design. *Journal of Educational Measurement*, 47, 286-298.

Table 18: Anchor item fit to blueprint – To February 2021

Area*	Actual	Target
10	5%	4%
20	14%	11%
30	10%	13%
40	14%	13%
50	10%	11%
60	10%	13%
70	14%	13%
80	12%	11%
90	12%	11%

*See Table 14 for the full name of each functional area.

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker equating was considered the preferred method.

Table 22 shows some of the parameters used to derive the equating estimates, along with other parameters describing the test forms. Of note is that on the anchor items, the candidates taking the October 2021 CHRP-KE scored modestly lower than the candidates taking the February 2021 CHRP-KE (68.3% vs. 70.1%; $t(529)=1.62$, *ns*). Because the October 2021 CHRP-KE candidates scored lower (non-significance notwithstanding), they would likely have a lower pass rate as compared to February 2021 candidates.

The equating analysis bears this out (Table 23). All methods indicate a pass mark of 95 or 96, with the preferred Tucker method providing a value of 96. The pass rate based on this equating run is lower, as expected, than what was seen in February 2021. The Tucker equating value of 95.15 was extracted from this analysis for use in setting the final pass mark.

Table 19: Equating parameter table – Total pass mark, to February 2021

		Feb. 2021	Oct. 2021
N		236	295
Scored items		151	151
Mean score	Total	70.5%	68.0%
	Anchors	70.1%	68.3%

Table 20: Equating outcome table – Total pass mark, to February 2021

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First-time
Equating Feb. 2021	96.59	97	67.4%	74.6%
Tucker	95.15	96	59.9%	68.5%
Levine observed	95.65	96	59.9%	68.5%
Mean	94.95	95	62.4%	70.5%
Circle Arc 1	95.07	96	59.9%	68.5%
Circle Arc 2	95.06	96	59.9%	68.5%

Equating Back to the August 2020 Administration

Linear equating was the chosen method for setting the pass mark. Linear equating is preferred with more than 100 candidates, and equipercentile equating is preferred with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc¹² equating is most prudent.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 SD units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form (and equivalent within each functional area) that adhered to the blueprint. Items with an increase or decrease of 10% in terms of difficulty were also removed as anchors. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.68 and a mean corrected point-biserial of .25 (for October 2021 candidates).

Table 21 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are well-aligned with targets and reflect the scope and approximate weighting across the full exam.

¹² Kim, S., & Livingston, S.A. (2010). Comparisons among small sample equating methods in a common-item design. *Journal of Educational Measurement*, 47, 286-298.

Table 21: Anchor item fit to blueprint – To August 2020

Area*	Actual	Target
10	4%	4%
20	11%	11%
30	11%	13%
40	13%	13%
50	11%	11%
60	13%	13%
70	13%	13%
80	13%	11%
90	11%	11%

*See Table 14 for the full name of each functional area.

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker equating was considered the preferred method. Note that given the difference in anchor item performance, Levin observed-score was also considered, though other metrics pointed to Tucker.

Table 22 shows some of the parameters used to derive the equating estimates, along with other parameters describing the test forms. Of note is that on the anchor items, the candidates taking the October 2021 CHRP-KE scored lower than the candidates taking the August 2020 CHRP-KE (68.4% vs. 72.0%; $t(534)=3.26$, $p<.01$). Because the October 2021 CHRP-KE candidates scored lower, they would likely have a lower pass rate as compared to August 2020 candidates.

The equating analysis bears this out (Table 23). All methods indicate a pass mark of 96 or 97, with the preferred Tucker method providing a value of 96. The pass rate based on this equating run is lower, as expected, than what was seen in August 2020. The Tucker equating value of 95.83 was extracted from this analysis for use in setting the final pass mark.

Table 22: Equating parameter table – Total pass mark, to August 2020

		Aug. 2020	Oct. 2021
N		241	295
Scored items		150	151
Mean score	Total	71.1%	68.0%
	Anchors	72.0%	68.4%

Table 23: Equating outcome table – Total pass mark, to August 2020

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First-time
Equating Aug. 2020	96.00	96	70.1%	75.9%
Tucker	95.83	96	59.9%	68.5%
Levine observed	96.83	97	57.4%	67.1%
Mean	96.22	97	57.4%	67.1%
Circle Arc 1	96.59	97	57.4%	67.1%
Circle Arc 2	96.59	97	57.4%	67.1%

Combined Results

Table 24 shows the pass mark values across the 3 equating runs. The value highlighted in green is the one that would be selected based on sample parameters at each equating run. The weighted mean (by number of anchor items and number of candidates) of the 3 identified values was the preliminary pass mark for the October 2021 CHRP-KE (94.93).

With a pass mark of 95, the pass rate for first-time October 2021 candidates was 70.5%, lower than seen in June 2021, February 2021, and August 2020 (as expected based on anchor set performance).

The processes used to derive the equated pass mark was presented to the CHRP EVC (see Table 26) via teleconference on November 12, 2021. The EVC voted unanimously to adopt the recommended pass mark. The HRP A Registrar approved the committee's recommendation and the pass mark was formally established.

Table 24: Equating outcome table – Combined results

	Jun. '21	Feb. '21	Aug. '20
Tucker	93.9	95.2	95.8
Levine observed	94.0	95.7	96.8
Mean	93.8	95.0	96.2
Circle arc 1	93.7	95.1	96.6
Circle arc 2	93.6	95.1	96.6

Table 25: Historical pass rates

	All	1st time
Feb. 18	64.2%	70.4%
Jun. 18	58.6%	66.2%
Oct. 18	67.0%	75.8%
Feb. 19	61.9%	72.5%
Jun. 19	56.6%	65.6%
Oct. 19	66.2%	74.3%
Feb. 20	65.3%	76.4%
Aug. 20	70.1%	75.9%
Feb. 21	67.4%	74.6%
Jun. 21	65.5%	72.7%
Oct. 21	62.4%	70.5%

Table 26: CHRP Examination Validation Committee members – Pass mark approval

Member	Credential	Years of Relevant Experience	Joined EVC	Industry
✓ Sunday Ajao	CHRL	15–20	2017	Banking/Finance
✓ Nancy Brandon	CHRL	20–25	2021	Power and Utilities
Roxanne Chartrand	CHRL	20–29	2018	Insurance
✓ Claire Chester	CHRL	10–15	2017	Long term care facility
Patrizia Finucan	CHRL	10–15	2021	Education
Tanya Gopaul	CHRL	10–15	2017	Banking
Annette Lawrence	CHRL	5–10	2021	Non-profit
Suman Seth	CHRL	15–19	2018	Public sector/education
Kriss Stone	CHRL	10–15	2017	
✓ Michelle Sultan	CHRL	10–15	2021	Education
Ileean Tait	CHRL	15–20	2017	Environmental
✓ Patricia Verkley	CHRL	10–15	2019	Not-for-profit
Karen Weiler	CHRL	20–29	2017	Software/ Communications

✓ Participated in the session.

Scoring

To finalize the scoring, repeat and outlier candidates who were not included in the item and form analysis were reinserted into the dataset. Scores for each of the 9 functional areas were also computed for each candidate. An Excel file with the final candidate results was provided to HRP.

Table 27 provides the means and standard deviations for the functional areas and for the total score, using all candidates who took the new October 2021 CHRP-KE forms. Table 28 provides the correlations between all functional areas. Caution should be exercised in interpreting differences between correlations. Variation can be explained largely by the number of items making up each functional area score. That is, functional areas with fewer items on the exam have lower correlations with the other functional areas. Figure 3 shows the distribution of scores for all candidates, along with the pass mark.

Table 27: Total and functional area scores for all candidates

Functional Area	Percentage	Mean	SD*
10 Strategy	66%	4.0	1.2
20 Professional Practice	70%	11.9	2.5
30 Organizational Effectiveness	64%	12.1	2.8
40 Workforce Planning & Talent Management	66%	12.5	2.7
50 Labour & Employee Relations	67%	11.3	2.3
60 Total Rewards	65%	12.3	2.8
70 Learning & Development	66%	13.1	3.0
80 Health, Wellness & Safe Workplace	68%	11.5	2.3
90 HR Metrics, Reporting & Financial Management	65%	11.1	2.6
Total score	66.1%	99.9	16.2

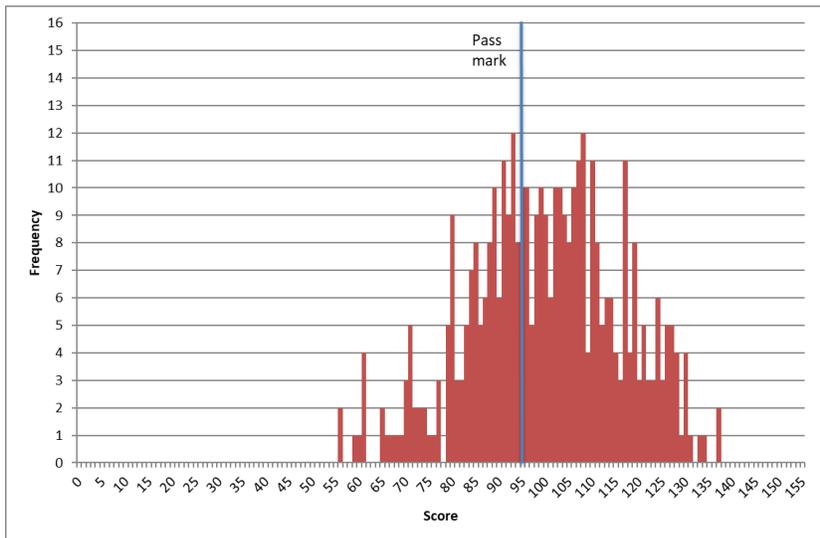
*SD = standard deviation.

Table 28: Correlations between functional area scores for all candidates

Area*	10	20	30	40	50	60	70	80	90
10		.30	.26	.32	.26	.33	.34	.26	.31
20			.52	.49	.41	.54	.50	.40	.42
30				.53	.46	.58	.55	.43	.45
40					.50	.55	.55	.43	.44
50						.46	.56	.39	.37
60							.61	.47	.51
70								.46	.43
80									.37
90									

*See Table 27 for the full name of each functional area.

Figure 3: Score distribution for all candidates



Key Examination Metrics

Table 29 shows the key examination metrics for candidates included in the main analysis; that is, only first-time candidates, with outliers removed. Past metrics are provided for reference.

Table 29: Key examination metrics – Candidates included in analysis only

Index	October 2021	June 2021	February 2021	August 2020	February 2020
Scored items	151	150	151	150	153
Candidates	295	267	236	241	144
Mean	102.7 (68.0%)	106.1 (70.7%)	106.5 (70.5%)	106.6 (71.1%)	109.1 (71.3%)
Median	104 (68.9%)	107 (71.3%)	108 (71.5%)	108 (72.0%)	110 (71.9%)
Skewness	-0.437	-0.607	-0.413	-0.402	-0.525
Kurtosis ⁱ	-0.161	0.632	-0.206	-0.498	-0.379
Range	56–137 (37.1– 90.7%)	51–141 (34.0– 94.0%)	56–138 (37.1– 91.4%)	67–133 (44.7– 88.7%)	66–137 (43.1– 89.5%)
Standard deviation	16.67	17.05	16.67	14.88	16.59
Cronbach's alpha	.90	.91	.90	.88	.90
Mean r_{pb}^*	.23	.24	.23	.20	.23
SEM ⁱⁱ	5.26	5.15	5.16	5.17	5.17
SEM at the pass mark	5.60	5.62	5.60	5.61	5.64
Decision consistency (uncorrected) ⁱⁱⁱ	.88	.88	.88	.89	.90
Perceived fairness ^{iv}	40%	47%	41%	45%	49%
Pass mark	94.928	96.017	96.590	96.000	98.458
Effective pass mark	95	97	97	96	99
Pass rate	70.5%	72.7%	74.6%	75.9%	76.4%

ⁱExcess

ⁱⁱSEM = standard error of measurement.

ⁱⁱⁱSubkoviak method.

^{iv}Based on responses to the post-examination survey. Value here may differ from that presented in main body of report because this value includes only candidates in the analysis.

Related Development Activities

Since the last administration of the CHRP-KE in June 2021, the following exam development activities have taken place.

Item Writing

To fill gaps in the bank and renew content, item writing was conducted in April–June 2021. Item writers (see Table 30) were identified by HRP A and trained in a remote session by Wickett on April 12 and 14, 2021. Items were written for both the CHRP-KE and CHRL-KE during this activity.

Table 30: Item writers

Writer	Credentials	Location	Years of HR Experience	Industry
Cal Barber	B. Comm, MBA	Toronto	40	Education
Nicole Bonenfant	CHRP, BA, MIRHR	Hamilton, Toronto	12	Health care, Education
Carolyn Capretta	MBA, CTPD	Hamilton	12	Education
Maria Ferraro	MBA, MSC, CGA, CPA	London	30	Education
Shady Hana	PhD, MBA	Toronto	13	Education, Consulting
Trina McGarvey	MIR, CHRL	Napanee	20	Government
Steve Risavy	PhD, MASc, CHRL, CHRP	Waterloo	11	Education
Julie Schermer	PhD	London	24	Education
Nicole Vincic	CHRL, BA, MA	Hamilton	22	Education
Gordon Wang	PhD	Toronto	17	Education

The item writers were provided with training via teleconference, and received additional materials covering the main elements of the training. The general guidance for writing quality multiple choice items was drawn primarily from Haladyna & Rodriguez (2013).¹³

Each item writer was selected based on expertise in identified functional areas, and they were assigned items within those functional areas. More specifically, each item writer was assigned competencies (drawn from the *HRPA Professional Competency Framework* [2014]) that were to be the focus of their items. Item writers were assigned 12 to 35 items each to write, for a total of 224 items.

¹³ Haladyna, T. M., & Rodriguez, M.C. (2013). *Developing and validating test items*. New York, NY: Routledge.

The item writers had access to the style guide that governs language usage on the HRP A exams and were provided with recent electronic textbooks as necessary. Item writers were required to include at least one authoritative source to back up each test item, and also provide rationales for the correct and incorrect answers.

Each item writer worked remotely, sending items to Wickett for review and comment via a secure file share site. Items were exchanged until such time as the item writer was comfortable with the content and Wickett was comfortable that the item would be successful at validation and upon use with candidates. This generally required several iterations per item.

Once all items were drafted and declared complete, they went through a reference verification to ensure the references were accurate and backed up each item. After that, they were sent to a certified professional editor for editorial. Items were adjusted based on this input and comments noted if future reviewers would need to attend to specific content concerns.

Appendix A

Blueprint

CHRP-Knowledge Examination

Human Resources Professionals Association
Version 2.2

Approved by CHRP Exam Validation Committee April 9, 2018

Approved by HRPA Registrar April 11, 2018

Effective June 2018

Credential

Passing the CHRP-Knowledge Examination is a requirement for certification for CHRP candidates. The examination reflects the *HRPA Professional HR Competency Framework* (2014).

Purpose

The CHRP-KE assesses whether a candidate has the level of discipline-specific knowledge necessary to practise human resources management at the CHRP level in a manner that is consistent with the protection of the public interest. Knowledge related exclusively to employment and workplace legislation is assessed on the CHRP Employment Law Examination.

Structure

The structural variables provide high-level guidance as to what the examination will be like.

Table 31: CHRP-KE Blueprint structural variables

Item types	Independent 4-option multiple choice
Length	175 items in total
	20–30 experimental items
Duration	Up to 3½ hours
Delivery mode	Computer-based testing in proctored test centres
Frequency	3 windows per year

Content Weighting

The functional area weights were set in 2014 to reflect an equal importance across the functional areas, except with a lower expectation for Strategy. The weights were modified slightly in 2018 to remove weighting for competencies most appropriately tested on the CHRP

Employment Law Examination. Within each functional area, items are distributed roughly evenly across the related competencies.

Table 32: Functional area weights on the CHRP-KE

Functional Area		Weight	Range
10	Strategy	4%	+/- 1%
20	Professional Practice	11%	+/- 2%
30	Organizational Effectiveness	13%	+/- 2%
40	Workforce Planning & Talent Management	13%	+/- 2%
50	Labour & Employee Relations	11%	+/- 2%
60	Total Rewards	13%	+/- 2%
70	Learning & Development	13%	+/- 2%
80	Health, Wellness & Safe Workplace	11%	+/- 2%
90	Human Resources Metrics, Reporting & Financial Management	11%	+/- 2%

Table 33: Competencies not eligible on the CHRP-KE

FA	Comp	FA	Comp	FA	Comp	FA	Comp
10	C005	40	C084	70	C152	80	C177
	C007		C089		C155		C179
	C009		50		C113		C156
	C011	C114			C158		C192
	C012	C117			C159	90	C194
	C017	C123			C163		C195
20	C035	C125			C165		C196
	C036	60	C139		C166		C204
	C037		C141		C171		C205
	C041		C143		C172		C206
30	C050		C146		C173	C210	
	C056		C175				
	C057						
	C065						

Minor amendments made November 20, 2018, by CHRP EVC, with approval of the Registrar.

Appendix B

MODIFIED ANGOFF METHOD

WHAT IT IS → The Modified Angoff method of setting cut scores is the most popular method used with high-stakes examinations. With this method, experts evaluate each item on a test for difficulty and judge how likely it is that someone who is borderline in performance will get each item correct. Borderline candidates have, by definition, just enough competence to be considered competent (e.g., to pass the test). Any candidate showing the same or a higher level of performance as a borderline candidate is thus a “passing” candidate, and any candidate showing performance below the level of a borderline candidate is a “failing” candidate. The method has been successfully defended in court as being a fair method of setting cut scores that are used to make high-stakes decisions about candidates.

HOW IT'S DONE → The Modified Angoff method typically requires 5 to 15 experts in the field and is facilitated by a psychometrician. There are many variations of the Modified Angoff method used in practice, but generally the process begins with detailed training on how to apply ratings, followed by development of a description of the borderline candidate. Once training is complete (including a calibration exercise to make sure all raters have fully grasped the method), ratings are applied individually by each rater and compiled by the psychometrician. Discrepancies across raters are identified and flagged for discussion. Raters then have an opportunity to discuss their ratings and to rerate any items if the new information is considered cause to do so. In some cases, the psychometrician will introduce data from previous administrations of the item to further refine judgments. Once all items have been rated, an average Angoff rating for the exam is calculated by simply taking the average of all item ratings. The result is the cut score for the exam as a whole.

WHY IT'S USED → The benefit of the Modified Angoff method is that the resulting cut scores set an objective hurdle for candidates. Candidates who demonstrate performance above the borderline level (as systematically established by experts) are considered to have sufficient competence, and those below that level are considered to have insufficient competence. The proportion of candidates deemed below or above the cut score is not arbitrary and depends only on the actual ability of those candidates. For examinations resulting in pass/fail decisions, the implication of this is that all candidates would pass if they all showed better than the minimal accepted level of competence (i.e., above the borderline), or they would all fail if they all showed less than the minimal accepted level of competence. What is important is whether each candidate scores above or below the cut score, with that cut score being set based on the actual difficulty of the test and the expected performance of candidates showing the lowest level of acceptable performance. Because of this, the Modified Angoff method fairly assesses individual candidates on their own merits.

References

- Cizek, G.J., & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Plake, B.S., & Cizek, G.J. (2012). Variations on a theme: The modified Angoff, extended Angoff, and yes/no standard setting methods. In G.J. Cizek (Ed.), *Setting performance standards* (pp. 181–199). New York, NY: Routledge.
- Smith, I.L., & Springer, C.C. (2009). Standard setting. In Institute for Credentialing Excellence, *Certification: The ICE handbook* (pp. 235–264). Washington, DC: Institute for Credentialing Excellence.



© 2015 Wickett Measurement Systems Inc.