PUBLIC RELEASE

Technical Report: February 2021 CHRP-KE

HRPA | Human Resources Professionals Association

7 March 2021

ICKETT
MEASUREMENT SYSTEMS

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 14
Key Validation 15
Establishing the Pass Mark: Equating 16
Scoring 22
Key Examination Metrics **25**
Related Development Activities **26**
Validation 26
Appendix A **28**
Appendix B **30**

 © 2021 Wickett Measurement Systems

List of Tables

Table 1: Test forms as administered 7
Table 2: Content-related post-examination survey questions* 11
Table 3: Administration-related post-examination survey questions* 12
Table 4: Testing location 12
Table 5: Preferred location (live remove proctoring candidates) 13
Table 6: Reason for choosing own location (live remove proctoring candidates) 13
Table 7: Evaluation of testing experience (live remove proctoring candidates) 13
Table 8: Value in future candidates being able to test from their own location (live remote proctoring candidates) 13
Table 9: Initial examination statistics – Combined across forms 14
Table 10: Section item statistics 14
Table 11: CHRP Examination Validation Committee members – Key validation 15
Table 12: Final scored examination fit to blueprint 16
Table 13: Anchor item fit to blueprint – To August 2020 18
Table 14: Equating parameter table – Total pass mark, to August 2020 18
Table 15: Equating outcome table – Total pass mark, to August 2020 19
Table 16: Anchor item fit to blueprint – To February 2020 20
Table 17: Equating parameter table – Total pass mark, to February 2020 20
Table 18: Equating outcome table – Total pass mark, to February 2020 21
Table 19: Equating outcome table – Combined results 21
Table 20: Historical pass rates 22
Table 21: CHRP Examination Validation Committee members – Pass mark approval 22
Table 22: Total and functional area scores for all candidates 23
Table 23: Correlations between functional area scores for all candidates 24
Table 24: Key examination metrics – Candidates included in analysis only 25
Table 25: CHRP Examination Validation Committee members – Validation 26
Table 26: CHRP-KE Blueprint structural variables 28
Table 27: Functional area weights on the CHRP-KE 29
Table 28: Competencies not eligible on the CHRP-KE 29

© 2021 Wickett Measurement Systems

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.....24

© 2021 Wickett Measurement Systems

Executive Summary[1]

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-KE[2]) was administered to 316 candidates using computer-based testing at Prometric test centres February 8–24, 2020, 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 candidates[3] (*n*=236) was 106.5 (70.5%), and for all candidates it was 103.0 (68.2%), 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 August 2020 and February 2020 administrations, yielding an integer pass mark of 97. 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 74.6% and a pass rate for all candidates of 67.4%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards[4] and ISO 17024 standards.[5]

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

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

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

© 2021 Wickett Measurement Systems

Administration

Form Setting

Using only validated test items, Wickett Measurement Systems prepared 3 new 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 year
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
- Item difficulty
- Item discrimination (adjusted point-biserial)
- Number of experimental items
- Adherence to blueprint
- Number of anchor items

The final form composition for the February 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, HRPA 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.

© 2021 Wickett Measurement Systems

Table 1: Test forms as administered

	Functional Area	Actual Items	Target
10	Strategy	7	6–8
20	Professional Practice	20	18–21
30	Organizational Effectiveness	22	21–24
40	Workforce Planning & Talent Management	22	21–24
50	Labour & Employee Relations	20	18–21
60	Total Rewards	22	21–24
70	Learning & Development	23	21–24
80	Health, Wellness & Safe Workplace	19	18–21
90	HR Metrics, Reporting & Financial Management	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 February 8–24, 2021, inclusive, and 316 candidates wrote the exam[6].

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. Standard security methods (as per Prometric protocols[7]) 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.

[6] Due to technical difficulties requiring the rescheduling of some candidates, testing continued through to February 26, 2021.

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

© 2021 Wickett Measurement Systems

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 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, 37 minutes (1 minute less than in August 2020; on average, form A candidates took 2 hours, 38 minutes, form B candidates took 2 hours, 39 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. Four candidates who were granted additional time as a testing accommodation exceeded the time limit of 3½ hours.

Seventeen candidates (5%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 5 candidates (1.6%) 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 −.07 for form A, negligible at a value of .00 for form B, and small at a value of −.05 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 a decrease of 1.6 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 clear pattern for candidates to prefer to book towards the end of the window rather than the start (though some of this effect is due to required rescheduling of candidates from earlier in the window to the end).

© 2021 Wickett Measurement Systems

Figure 1: Examination time distribution for all candidates

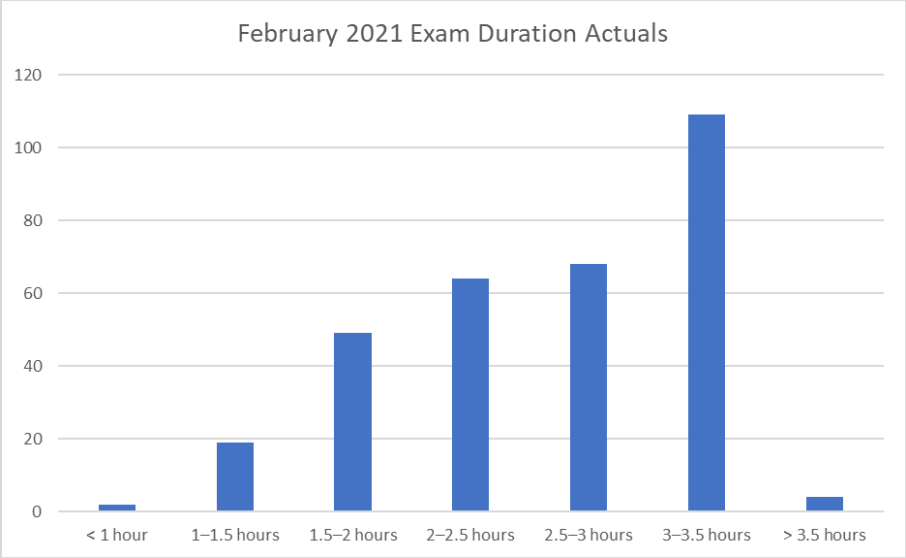
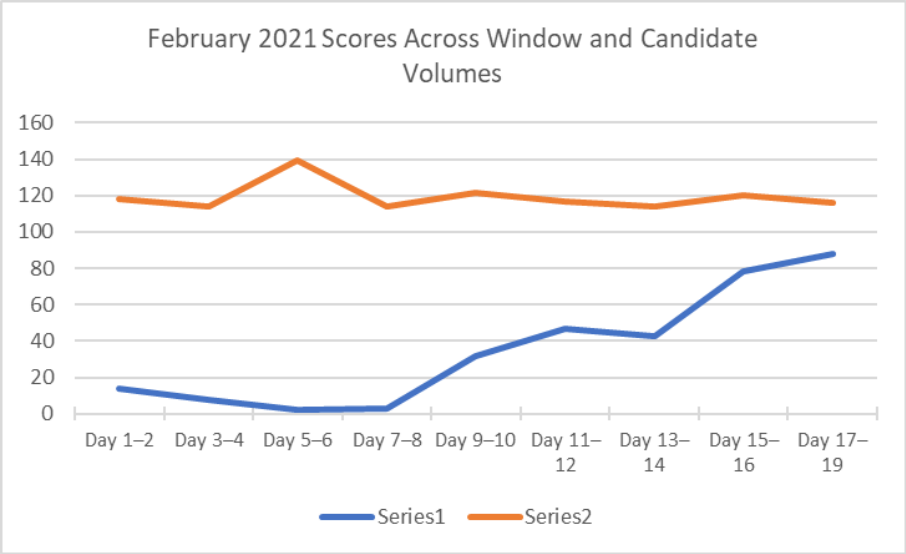


Figure 2: Candidate volume and score trends across testing window



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. Three 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. All outliers were removed from initial analyses; candidates with abnormal response patterns were also removed. 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, 3 candidates were removed from analysis.

 © 2021 Wickett Measurement Systems

Candidates who had failed a previous HRPA examination (CKE, CHRP-KE, or CKE 2) scored lower than did those who had not (61.4% and 69.2%, respectively, on the full exam of 175 items). This difference was meaningful and significant ($t(176)=6.54$, $p<.001$). In keeping with standard procedures, these candidates were removed from subsequent analyses. The CHRP-KE analysis proceeded with 236 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; 314 responses were obtained from candidates (response rate, 99%). The survey was revised as of the August 2020 administration to collect information specifically applicable to the use of live remote proctoring.

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.

© 2021 Wickett Measurement Systems

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.	176	110	12	11	5	4.4	91%	91%
2.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	93	131	42	33	12	4.0	72%	67%
3.	I feel I was adequately prepared to write this examination.	24	120	106	54	7	3.7	46%	52%
4.	The questions in the examination were clearly written.	29	167	67	47	2	3.8	63%	63%
5.	The terminology used in the examination was accurate.	38	172	77	24	0	4.0	68%	72%
6.	The situations presented in the examination were realistic.	40	195	62	12	1	4.0	76%	81%
7.	The questions in the examination reflected the examination blueprint.	30	120	109	45	7	3.7	48%	57%
8.	The examination was a fair assessment of my ability.	21	102	95	65	26	3.4	40%	48%

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

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

© 2021 Wickett Measurement Systems

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.	146	119	15	20	9	4.3	86%	91%
10.	I was well informed about the examination rules and regulations.	166	130	3	6	3	4.5	96%	96%
11.	Proctors enforced the exam-day rules.	184	112	8	1	3	4.6	96%	98%
12.	Proctors were professional and courteous.	193	97	8	3	7	4.5	94%	96%
13.	The tutorial helped me understand how to complete the examination on the computer.	138	137	19	11	2	4.4	90%	95%
14.	Navigation through the examination was easy and intuitive.	178	109	9	9	4	4.5	93%	97%

*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.

Table 4: Testing location

Response	Count	%
Test centre	2	1%
Own location	307	99%

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 HRPA 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.

 © 2021 Wickett Measurement Systems

Table 5: Preferred location (live remove proctoring candidates)

Response	Count	%
I preferred using my own location.	185	60%
I preferred going to a test centre.	68	22%
I have no preference.	54	18%

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

Response	Count	%
No test centres were open in my area.	123	40%
I preferred to avoid being around other people.	36	12%
I liked the convenience of not having to travel to a test centre.	71	23%
I felt like I would perform better in my own environment.	27	9%
Other (please specify)	50	16%

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

	Count	%
Very positive	92	30%
Positive	127	42%
Neutral	42	14%
Negative	26	8%
Very negative	19	6%

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

Response	Count	%
Yes	292	95%
No	14	5%

Because only 2 candidates who wrote in a test centre had post-exam survey data, table are not shown.

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 HRPA for information and consideration. Nothing actionable with respect to scoring emerged in these comments.

 © 2021 Wickett Measurement Systems

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 9. The section statistics are presented in Table 10.

Table 9: Initial examination statistics – Combined across forms

Index	Value
Items	151
Total candidates	316
Candidates in analysis	236
Mean score	106.5 (70.5%)
Score range	56–138 (37.1–91.4%)
Cronbach's alpha	.90
Mean r_{pb}*	.23

Table 10: Section item statistics

Index	Section 1	Section 2
Total items	88	87
Scored items	76	75
Candidates in analysis	236	
Mean	53.6 (70.5%)	52.9 (70.6%)
Standard deviation	8.6	8.8
Range	29–70	26–73

Because so few candidates wrote in a test centre, analyses of potential performance differences were not conducted.

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

 © 2021 Wickett Measurement Systems

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 the 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 March 1, 2021, using members of the CHRP Examination Validation Committee (EVC). The EVC (Table 11) 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 11: CHRP Examination Validation Committee members – Key validation

Member	Credential	Years of Relevant Experience	Joined EVC	Industry
Sunday Ajao	CHRL	15–20	2017	Banking/finance
Roxanne Chartrand	CHRL	20–29	2018	Insurance
✓ **Claire Chester**	CHRL	10–15	2017	Health services
✓ **Tanya Gopaul**	CHRL	10–15	2017	Banking
✓ **Jean Lazarus**	CHRL	15–19	2017	Health services
Suman Seth	CHRL	15–19	2018	Government
Kriss Stone	CHRL	10–15	2017	Real estate
✓ **Ielean 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 group 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 group was informed that no items fell below the flagging criteria, and as such all were considered appropriate to use in making decisions on candidates. The group endorsed this

 © 2021 Wickett Measurement Systems

recommendation and the set of 151 items were approved for use in scoring the February 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 35.2% to 94.9%, with a mean of 70.5%. The r_{pb}* values ranged from .03 to .48, with a mean of .23.

Table 12 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 12: 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),[8] was used to establish the pass mark for the February 2021 CHRP-KE. The goal of this process was to set a pass mark for the February 2021 CHRP-KE that would be equivalent to that set for previous CHRP-KE administrations; that is, to set a pass mark that would give each candidate the same probability of passing regardless of which form they took.

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

 © 2021 Wickett Measurement Systems

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.

Two equating procedures were conducted back to different administrations (August 2020 and February 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 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 arc[9] 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 0.71 and a mean corrected point-biserial of .24 (for February 2021 candidates).

Table 13 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.

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

 © 2021 Wickett Measurement Systems

Table 13: Anchor item fit to blueprint – To August 2020

Area*	Actual	Target
10	3%	4%
20	13%	11%
30	13%	13%
40	15%	13%
50	8%	11%
60	13%	13%
70	13%	13%
80	13%	11%
90	13%	11%

*See Table 12 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 14 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 February 2021 CHRP-KE scored negligibly lower than the candidates taking the August 2020 CHRP-KE (70.6% vs. 71.3%; $t(475)=0.58$, *ns*). Because the February 2021 CHRP-KE candidates scored lower (non-significance notwithstanding), they would likely have a slightly lower pass rate as compared to August 2020 candidates.

The equating analysis bears this out (Table 15). 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 just slightly lower, as expected, than what was seen in August 2020. The Tucker equating value of 95.66 was extracted from this analysis for use in setting the final pass mark.

Table 14: Equating parameter table – Total pass mark, to August 2020

		Aug. 2020	Feb. 2021
	N	241	236
	Scored items	150	151
Mean score	Total	71.1%	70.5%
	Anchors	71.3%	70.6%

 © 2021 Wickett Measurement Systems

Table 15: Equating outcome table – Total pass mark, to August 2020

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First-time**
Equating Aug. 2020	**96.00**	**96**	**70.1%**	**75.9%**
Tucker	95.66	96	*69.0%*	*75.4%*
Levine observed	95.97	96	*69.0%*	*75.4%*
Mean	96.65	97	*67.4%*	*74.6%*
Circle Arc 1	96.69	97	*67.4%*	*74.6%*
Circle Arc 2	96.69	97	*67.4%*	*74.6%*

Equating Back to the February 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 arc[10] 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 0.70 and a mean corrected point-biserial of .23 (for February 2021 candidates).

Table 16 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.

 © 2021 Wickett Measurement Systems

Table 16: Anchor item fit to blueprint – To February 2020

Area*	Actual	Target
10	6%	4%
20	13%	11%
30	13%	13%
40	15%	13%
50	8%	11%
60	10%	13%
70	10%	13%
80	13%	11%
90	13%	11%

*See Table 12 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 17 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 February 2021 CHRP-KE scored modestly lower than the candidates taking the February 2020 CHRP-KE (70.4% vs. 72.0%; $t(378)=1.26$, *ns*). Because the February 2021 CHRP-KE candidates scored modestly lower (non-significance notwithstanding), they would likely have a modestly lower pass rate as compared to February 2020 candidates.

The equating analysis bears this out (Table 18). All methods indicate a pass mark of 98 or 99, with the preferred Tucker method showing a 98. The pass rate based on this equating run is lower, as expected, than what was seen in February 2020. The Tucker equating value of 97.69 was extracted from this analysis for use in setting the final pass mark.

Table 17: Equating parameter table – Total pass mark, to February 2020

		Feb. 2020	Feb. 2021
	N	144	236
	Scored items	153	151
Mean score	Total	71.3%	70.5%
	Anchors	72.0%	70.4%

© 2021 Wickett Measurement Systems

Table 18: Equating outcome table – Total pass mark, to February 2020

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First-time**
Equating Feb. 2020	**98.46**	**99**	**65.3%**	**76.4%**
Tucker	97.69	98	*65.2%*	*72.9%*
Levine observed	98.04	99	*60.4%*	*67.8%*
Mean	97.87	98	*65.2%*	*72.9%*
Circle Arc 1	98.33	99	*60.4%*	*67.8%*
Circle Arc 2	98.33	99	*60.4%*	*67.8%*

Combined Results

Table 19 shows the pass mark values across the 2 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 2 identified values was the preliminary pass mark for the February 2021 CHRP-KE (96.59).

With a pass mark of 97, the pass rate for first-time February 2021 candidates was 74.6%, slightly lower than seen in August 2020 and February 2020 (as expected based on anchor set performance).

The processes used to derive the equated pass mark was presented to the CHRP EVC (

Table 21) via teleconference on March 4, 2021. The EVC voted unanimously to adopt the recommended pass mark. The HRPA Registrar later approved the committee's recommendation; the pass mark was formally established.

Table 19: Equating outcome table – Combined results

	Aug. '20	Feb. '20
Tucker	95.7	97.7
Levine observed	96.0	98.0
Mean	96.7	97.9
Circle arc 1	96.7	98.3
Circle arc 2	96.7	98.3

© 2021 Wickett Measurement Systems

Table 20: 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%

Table 21: 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
✓ **Roxanne Chartrand**	CHRL	20–29	2018	Insurance
Claire Chester	CHRL	10–15	2017	Health services
Tanya Gopaul	CHRL	10–15	2017	Banking
Jean Lazarus	CHRL	15–19	2017	Health services
Suman Seth	CHRL	15–19	2018	Government
✓ **Kriss Stone**	CHRL	10–15	2017	Real estate
Ielean 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 HRPA.

© 2021 Wickett Measurement Systems

Table 22 provides the means and standard deviations for the functional areas and for the total score, using all candidates who took the new February 2021 CHRP-KE forms. Table 23 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 22: Total and functional area scores for all candidates

	Functional Area	Percentage	Mean	SD*
10	**Strategy**	71%	4.2	1.3
20	**Professional Practice**	67%	11.3	2.5
30	**Organizational Effectiveness**	66%	12.6	2.9
40	**Workforce Planning & Talent Management**	70%	13.2	2.7
50	**Labour & Employee Relations**	70%	11.9	2.5
60	**Total Rewards**	67%	12.8	2.9
70	**Learning & Development**	70%	14.1	3.3
80	**Health, Wellness & Safe Workplace**	66%	11.2	2.6
90	**HR Metrics, Reporting & Financial Management**	68%	11.6	2.3
Total score		68.2%	103.0	17.4

*SD = standard deviation.

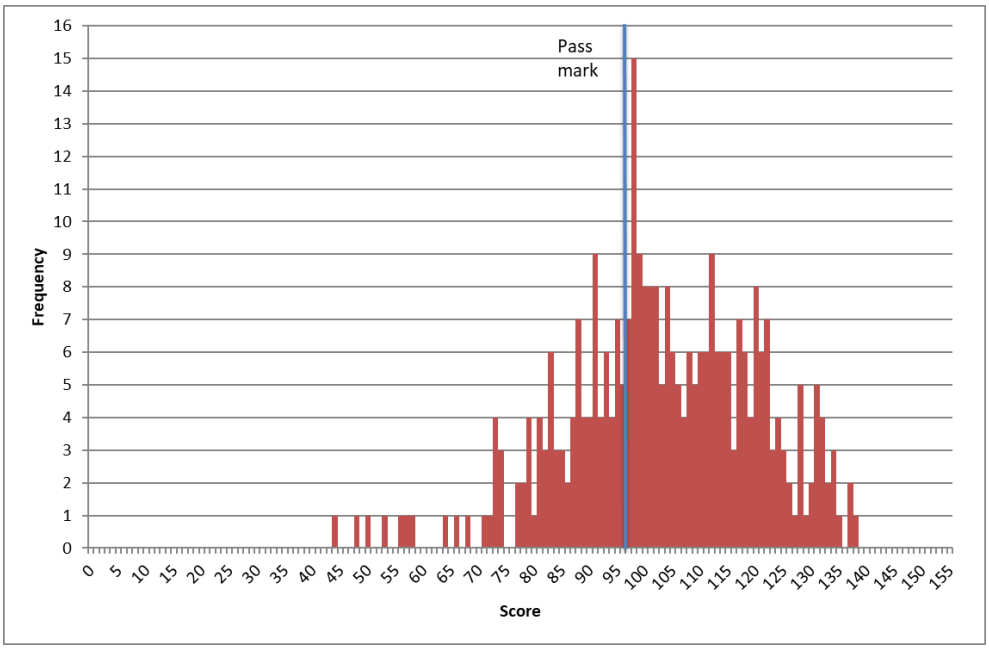
© 2021 Wickett Measurement Systems

Table 23: Correlations between functional area scores for all candidates

Area*	10	20	30	40	50	60	70	80	90
10		.27	.33	.40	.26	.38	.40	.38	.35
20			.54	.58	.56	.56	.55	.57	.48
30				.53	.49	.55	.59	.48	.47
40					.55	.59	.60	.56	.58
50						.47	.53	.55	.50
60							.60	.54	.53
70								.55	.50
80									.52
90									

*See Table 22 for the full name of each functional area.

Figure 3: Score distribution for all candidates



 © 2021 Wickett Measurement Systems

Key Examination Metrics

Table 24 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 24: Key examination metrics – Candidates included in analysis only

Index	February 2021	August 2020	February 2020	October 2019	June 2019
Scored items	151	150	153	153	150
Candidates	236	241	144	175	128
Mean	106.5 (70.5%)	106.6 (71.1%)	109.1 (71.3%)	106.4 (69.5%)	101.1 (67.4%)
Median	108 (71.5%)	108 (72.0%)	110 (71.9%)	107 (69.9%)	100.5 (67.0%)
Skewness	−0.413	−0.402	−0.525	−0.466	0.002
Kurtosis[i]	−0.206	−0.498	−0.379	−0.057	−0.446
Range	56–138 (37.1–91.4%)	67–133 (44.7–88.7%)	66–137 (43.1–89.5%)	61–138 (39.9–90.2%)	60–138 (40.0–92.0%)
Standard deviation	16.67	14.88	16.59	15.76	16.43
Cronbach's alpha	.90	.88	.90	.89	.90
Mean r_{pb}*	.23	.20	.23	.22	.22
SEM[ii]	5.16	5.17	5.17	5.21	5.23
SEM at the pass mark	5.60	5.61	5.64	5.59	5.57
Decision consistency (uncorrected)[iii]	.88	.89	.90	.88	.86
Perceived fairness[iv]	41%	45%	49%	53%	42%
Pass mark	96.590	96.000	98.458	97.499	93.247
Effective pass mark	97	96	99	98	94
Pass rate	74.6%	75.9%	76.4%	74.3%	65.6%

[i]Excess

[ii]SEM = standard error of measurement.

[iii]Subkoviac 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.

© 2021 Wickett Measurement Systems

Related Development Activities

Since the last administration of the CHRP-KE in August 2020, the following exam development activities have taken place.

Validation

To provide sufficient scorable items for upcoming administrations, a validation session was held with the EVC (seeTable 25) remotely on November 16, 18, 20 and 23, 2020. During these 4 sessions, CHRP ELE items were also validated, and that portion of the validation activity is reported in the technical report for that examination.

Note that scheduling precluded all identified members from being available for all 4 days; those marked as having participated attended the validation activity on at least 1 day.

Table 25: CHRP Examination Validation Committee members – Validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓ **Sunday Ajao**	CHRL	15–20	2017	Banking/Finance
✓ **Roxanne Chartrand**	CHRL	20–29	2018	Insurance
✓ **Claire Chester**	CHRL	10–14	2017	Regulation/CPA
✓ **Tanya Gopaul**	CHRL	10–15	2017	Banking
✓ **Jean Lazarus**	CHRL	15–19	2017	Health services
Suman Seth	CHRL	15–19	2018	Public sector
✓ **Kriss Stone**	CHRP	10–15	2017	Real estate
✓ **Ielean 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 EVC members received advance materials outlining:

- Purpose of the session
- Description of the CHRP credential
- CHRP-KE blueprint
- Criteria for good test items
- Validation process

 © 2021 Wickett Measurement Systems

The committee members received refresh training on the validation activity on the first day of the session. For participants not able to join on the first day, the received individual training on the first day of their involvement. Each day, committee members were provided with approximately 50–70 items via a secure file share site, and then worked individually reviewing items through the day, submitting their appraisal and any suggested revisions to Wickett through the day. They were directed to make sure the items reflected current practice and were suitable to make decisions about who should receive the CHRP credential.

At the end of each day, the committee convened online and were shown items flagged for revision. Where committee members proposed changes, these were discussed by the group before implementation.

For each item, the committee was asked to either:

- Validate the item for use in the next 3 years to make decisions about who would be certified as a CHRP
- Move the item to the CHRL-KE or CHRP ELE bank
- Revise the item to make it suitable for use
- Declare the item unsound and send it back for revision or removal from the bank

At the close of each day, committee members were walked through a process to verify deletion of all item files in use that day.

The committee validated 135 items as suitable for the CHRP-KE and rejected 4 items. Twelve items were revised prior to validation as part of this exercise. The committee also verified the functional area and competency for all items, and added rationales and references where missing, incomplete, or not current.

© 2021 Wickett Measurement Systems

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 26: 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

 © 2021 Wickett Measurement Systems

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

Table 27: 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 28: Competencies not eligible on the CHRP-KE

FA	Comp
10	C005
	C007
	C009
	C011
	C012
	C017
20	C035
	C036
	C037
	C041
30	C050
	C056
	C057
	C065

FA	Comp
40	C084
	C089
50	C113
	C114
	C117
	C123
	C125
60	C139
	C141
	C143
	C146

FA	Comp
70	C152
	C155
	C156
	C158
	C159
	C163
	C165
	C166
	C171
	C172
	C173
	C175

FA	Comp
80	C177
	C179
	C187
	C192
90	C194
	C195
	C196
	C204
	C205
	C206
	C210

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

© 2021 Wickett Measurement Systems

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.