

Technical Report: January 2019 CHRP ELE

Human Resources Professionals Association

28 January 2019



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.....	13
Establishing the Pass Mark: Equating	15
Scoring	24
Key Examination Metrics	26
Related Development Activities	27
Item Validation	27
Item Writing	28
Item Review	29
Appendix	31

List of Tables

Table 1: Domain fit at administration	6
Table 2: Cognitive level fit at administration	7
Table 3: Administration-related post-examination survey questions*	11
Table 4: Content-related post-examination survey questions*	11
Table 5: Preference regarding independent and case-based items.....	12
Table 6: Preference regarding number of response options	12
Table 7: Preference regarding computer-based testing versus pencil-and-paper	12
Table 8: Value of access to legislation	12
Table 9: Initial examination statistics	13
Table 10: CHRP Examination Validation Committee – Key validation	14
Table 11: Domain fit for final scored items	15
Table 12: Cognitive level fit for final scored items.....	15
Table 13: Item type fit for final scored items	15
Table 14: Anchor item fit to blueprint – To September 2018.....	17
Table 15: Equating parameter table – To September 2018	17
Table 16: Equating outcome table – To September 2018.....	18
Table 17: Anchor item fit to blueprint – To January 2018	19
Table 18: Equating parameter table – To January 2018.....	19
Table 19: Equating outcome table – To January 2018	20
Table 20: Anchor item fit to blueprint – To May 2018	21
Table 21: Equating parameter table – To May 2018.....	21
Table 22: Equating outcome table – To May 2018	22
Table 23: Equating outcome table – Combined results	23
Table 24: Historical pass rates	23
Table 25: CHRP Examination Validation Committee – Pass mark approval.....	24
Table 26: Total and domain scores for all candidates.....	24
Table 27: Correlations between functional area scores for all candidates	25
Table 28: Key examination metrics – Candidates included in analysis only.....	26
Table 29: CHRP Examination Validation Committee – Item validation	27
Table 30: Item writers.....	28
Table 31: Panel for item review session.....	30
Table 32: CHRP Employment Law Examination Blueprint Structural Variables.....	31

Table 33: CHRP Employment Law Examination Blueprint Content Weights33
Table 34: CHRP Employment Law Examination Blueprint Cognitive Level Weights.....34

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.....25

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 Employment Law Exam (CHRP ELE) was administered to 132 candidates using computer-based testing at Prometric test centres January 7–21, 2019, inclusive. The examination comprised 110 three-option multiple choice items and had a 3½-hour time limit.

As per the CHRP ELE blueprint, the exam was scored using the 98–102 best-performing items (while adhering to the prescribed distribution across topics). The mean score for first-time candidates ($n=119^2$) was 72.3 (71.6%), and for all candidates it was 71.6 (70.9%), out of 101 validated items for scoring. Reliability was marginal at .74 (noting that there is substantial range restriction with these candidates). The final set of scored items adhered to the blueprint parameters.

The pass mark was set using equating back to the January and May 2018 administrations, yielding an integer pass mark of 57. 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 ELE. This pass mark resulted in a pass rate for first-time candidates of 98.3% and a pass rate for all candidates of 97.0%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards³ and ISO 17024 standards.⁴

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

³ National Commission for Certifying Agencies (2014). *Standards for the accreditation of certification programs*. Washington, DC: Institute for Credentialing Excellence.

⁴ 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 one 110-item test form. Wickett constructed the final test form according to the following parameters:

1. Including only items validated by the validation panel in the past 2 years
2. Fitting the total item count of 110
3. Excluding enemy items
4. Matching the blueprint weights
5. Maximizing spread across subtopics as per the blueprint weights
6. Reducing item exposure
7. Selecting items with perceived psychometric effectiveness, using statistics from previous administrations as available

The final form was reviewed for currency and enemy items by Alyssa Young and Ilean Tait (CHRP Examination Validation Committee members) in a remote session held November 12, 2018.

The final form composition for the January 2018 CHRP ELE is shown in Table 1 (domain weighting) and Table 2 (cognitive level weighting). The form reflected the examination blueprint (see Appendix for full CHRP ELE blueprint).

Note that at any administration, HRPAs make 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: Domain fit at administration

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	48	46% ± 5%	46–56	—
B Employer Obligations	39	33% ± 4%	32–40	—
C Regulations and Legislation	23	21% ± 3%	20–26	—
TOTAL	110		110	—

Table 2: Cognitive level fit at administration

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	10	10% ± 3%	8–14	—
Application	65	60% ± 10%	55–77	—
Critical thinking	35	30% ± 10%	22–44	—
TOTAL	110		110	—

The test form adhered to the blueprint for content domain and cognitive level.

Testing Window

The examination was administered via computer-based testing at Prometric test sites primarily in Ontario. The testing window was January 7–21, 2019, inclusive, and 132 candidates wrote the exam. Note that one candidate wrote a large-font paper version as an accommodation.

Candidates had access to a basic-function calculator on screen and access via PDF to 10 pieces of searchable legislation (compiled into 2 documents):

Provincial

- AODA – *Accessibility for Ontarians with Disabilities Act, 2005*
- ESA – *Employment Standards Act, 2000*
- LRA – *Labour Relations Act, 1995*
- OHRC – *Human Rights Code*
- OHSA – *Occupational Health and Safety Act*
- PEA – *Pay Equity Act*
- WSIA – *Workplace Safety and Insurance Act, 1997*

Federal

- CHRA – *Canadian Human Rights Act*
- CLC – *Canada Labour Code*
- PIPEDA – *Personal Information Protection and Electronic Documents Act*

The versions of the legislation were as accessed on October 19, 2018.

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 110 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, 55 minutes (up 1 minute from September 2018). The time limit on the CHRP ELE was 3½ hours, suggesting that time may have been a factor for at least some candidates.

Sixteen candidates (12%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 8 candidates (6%) left at least one 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 presently they do not appear problematically high, though they are not as low as would be preferred. Compared with the September 2018 administration, there was little change in observed values. Note that because they have access to legislation, candidates may take more time than intended by researching more answers. This may skew time metrics higher.

The correlation between scores on the 110 items and time spent writing the examination was essentially zero at a value of $-.07$, suggesting that time constraints were not generally related to candidate performance. (Note that two candidates exceeded the time limit; these candidates were granted additional time in advance of the administration as an accommodation.)

Candidate scores were computed across the window to look for any evidence of item exposure. As shown in Figure 2, there was little variation across the window, and the difference between the first 3 days and the last 3 days was a modest decrease of 2.2 marks out of 110 (though there were few candidates overall so this analysis lacks power to identify a significant change).

As a matter of interest, candidate volumes were also examined across the window; these are also shown in Figure 2. The usual peak in volumes towards the end of the window was observed.

Figure 1: Examination time distribution for all candidates

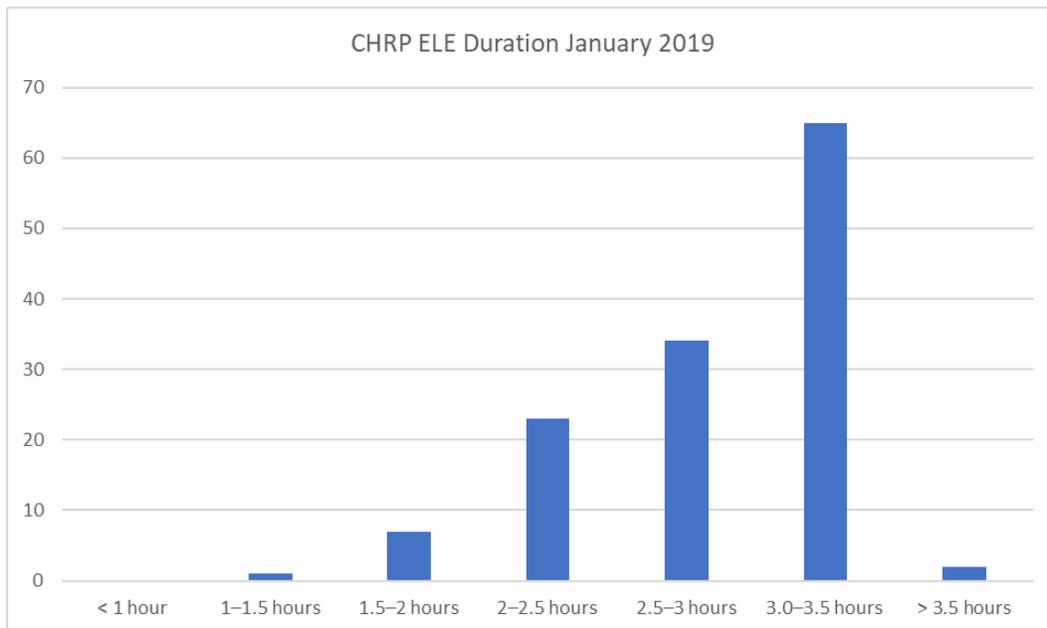
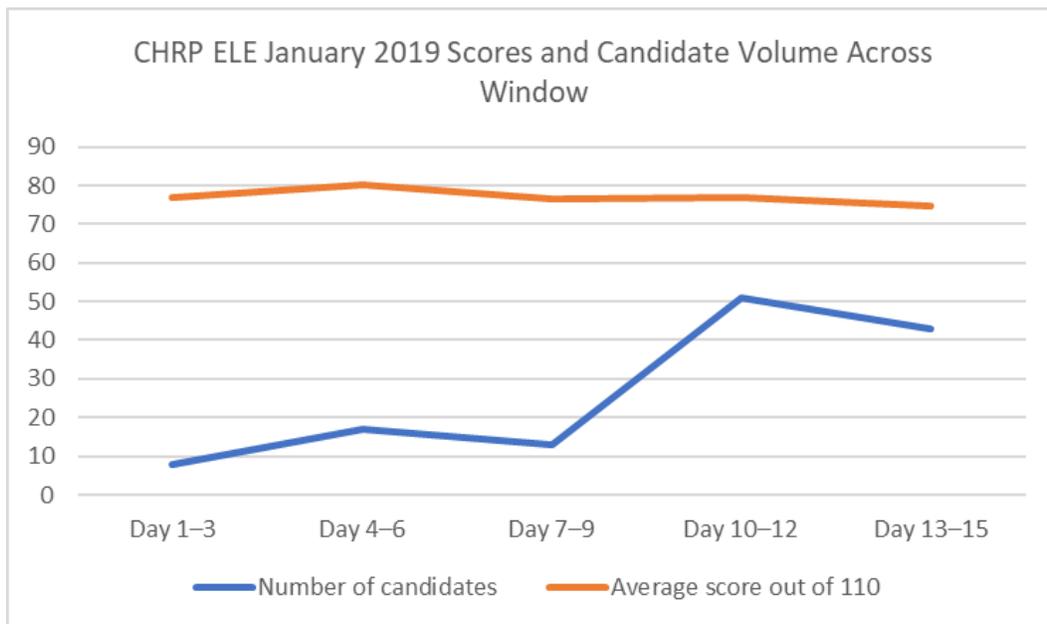


Figure 2: Candidate volume and score trends across testing window



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 110 items. Zero candidates were flagged for an abnormally low or high score (z value outside +/- 3.0). Also, the 110 items were arbitrarily broken into 4 blocks of 25 items for each candidate plus 1 final block of 10 items; the 5 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. No instances were identified and so no candidates were removed from analysis. Candidates with abnormal response patterns (such as having 5 or more blanks) were removed. To be conservative, candidates who had been granted a testing accommodation were also removed from the main analysis (simply because their testing conditions were not the same as those of the main group of candidates, even though each accommodation was granted on the premise that it would make the testing experience equivalent in terms of opportunity to demonstrate competence). As a result of these factors, 3 candidates were removed from analyses.

Candidates who had failed a previous employment law examination (CHRP ELE or CHRL ELE; $n=10$) scored lower than did those who had not (71.4 and 77.6, respectively, on the full exam of 110 items; $t(11)=2.81$, $p<.05$). In keeping with standard procedures, these candidates were removed from subsequent analyses. The CHRP ELE analysis proceeded with 119 candidates.

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

Post-Examination Survey

Candidates were provided access to the post-examination survey immediately after submitting their responses to the CHRP ELE; 128 candidates responded (response rate, 97%).

Table 3 shows the responses to the administration-related questions. Note that candidates were generally positive about the administration experience. Table 4 shows the content-related questions; there was a tendency to more neutrality on these questions.

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

	Question	SA	A	N	D	SD	Score	Agreement
1.	I was able to book a seat to write the examination at a time that was convenient for me.	68	44	5	10	1	4.31	88%
2.	I was well informed about what documents to bring to the exam location.	90	36	1	1	0	4.68	98%
3.	Proctors enforced the exam-day rules and the security procedures at the test centre were what I expected.	94	31	1	0	0	4.74	99%
4.	Proctors were professional and courteous.	93	27	3	1	1	4.68	96%
5.	The tutorial helped me understand how to complete the examination on the computer.	84	38	2	1	0	4.64	98%
6.	The legislation and case texts were easy to access during the examination.	66	48	4	7	0	4.38	91%
7.	Navigation through the examination was easy and intuitive.	79	39	6	1	0	4.57	94%

*Response categories: SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.

Table 4: Content-related post-examination survey questions*

	Question	SA	A	N	D	SD	Score	Agreement
8.	The time allotted for this examination was sufficient.	59	44	5	14	2	4.16	83%
9.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	48	60	7	7	2	4.17	87%
10.	I feel I was adequately prepared to write this examination.	17	69	26	8	1	3.77	71%
11.	The questions in the examination were clearly written.	16	70	18	16	1	3.69	71%
12.	The terminology used in the examination was accurate.	18	86	14	3	0	3.98	86%
13.	The situations presented in the examination were realistic.	24	84	9	3	0	4.08	90%
14.	The questions in the examination reflected the Employment Law Examination blueprint.	24	73	18	5	0	3.97	81%
15.	The examination was a fair assessment of my ability.	16	59	31	12	2	3.63	63%

*Response categories: SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.

Candidates were asked their opinions regarding several structural variables; these results appear in Table 5 through Table 8. Candidates reported a modest preference for independent items and a strong preference for 3-option multiple choice items. Most indicated that taking the test on a computer likely had no effect on their performance, and where a preference was indicated it was in favour of a benefit to completing the examination on a computer. Most indicated that access to the legislation was necessary, whether they consulted it a few times or often.

Table 5: Preference regarding independent and case-based items

	Count	%
I preferred the independent items.	50	42%
I preferred the case-based items.	37	31%
I had no preference between independent and case-based items.	33	28%

Table 6: Preference regarding number of response options

	Count	%
I preferred having 3 options.	87	73%
I preferred having 4 options.	7	6%
It did not matter to me how many options were used.	26	22%

Table 7: Preference regarding computer-based testing versus pencil-and-paper

Question	Count	%
I feel that completing the examination on a computer improved my performance.	39	33%
I feel that completing the examination on a computer decreased my performance.	8	7%
I feel that completing the examination on a computer had no effect on my performance.	73	61%

Table 8: Value of access to legislation

	Count	%
Yes, it was essential to me in completing the examination.	49	41%
Yes, but I only consulted it a few times.	57	48%
No, I could not find the answers to questions I had.	10	8%
No, I did not need to consult it to complete the examination.	2	2%
No, it was more of a distraction than an aid.	2	2%

An open-ended question was also posed to candidates asking for any additional comments. 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 ELE examination was 110 items, of which approximately 100 were to be scored. The remain 8–12 items were designated as experimental. However, because only 1 new form was administered, all items were potentially available for scoring and the focus of subsequent item analysis and key validation was on determining the best set of approximately 100 items that still reflected the examination blueprint.

The initial analysis summary statistics are presented in Table 9.

Table 9: Initial examination statistics

Index	CHRP ELE
Items	110
Total candidates	132
Candidates in analysis	119
Mean	77.8 (70.7%)
Standard deviation	7.92
Range	56–97 (50.9–88.2%)
Cronbach's alpha	.71
Disattenuated alpha	.85
Mean r_{pb}^*	.13

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 flagged 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, after removing items with difficulty values at the extremes. Items were ranked from worst performing to best performing accordingly.

Key Validation

Key validation was conducted via web meeting on January 25, 2019, using members of the CHRP Examination Validation Committee (EVC). The group (Table 10) was first reminded of the

methods used for key validation and was oriented to the main statistics used to evaluate the quality of the CHRP ELE.

Table 10: CHRP Examination Validation Committee – Key validation

Member	Credential	Years of Relevant Experience	Years on EVC	Industry
✓ Sunday Ajao	CHRL	15–20	2	Banking/Finance
✓ Roxanne Chartrand	CHRL	20–29	1	Insurance
Claire Chester	CHRL	10–14	2	Regulation/CPA
Tanya Gopaul	CHRL	10–15	2	Banking
Jean Lazarus	CHRL	15–19	2	Health services
Suman Seth	CHRL	15–19	1	Public sector
Kriss Stone	CHRP	10–15	2	Real estate
✓ Ielean Tait	CHRL	15–20	2	Environmental
Karen Weiler	CHRL	20–29	2	Software/ Communications
✓ Alyssa Young	CHRL	5–9	2	Non-profit

✓ Participated in the session.

The group was informed that test reliability, as measured by Cronbach's alpha, was .71 based on the set of 110 potentially scored items and that this was below the generally accepted threshold of .80. The group was advised that restriction of range was considered the most likely basis for the lower value. They were also informed that part of the goal of the key validation review was to bring this value up if possible.

The group was walked through the flagged items one at a time, with the recommendation that the worst-performing items be removed from scoring, but were given less direction on those with borderline statistics. Where available, candidates' comments about the items were also shown. The group made decisions based on content and the data through discussion; they removed 9 items that they felt were least appropriate to retain for scoring. Panel members' comments about specific items were recorded for future item revision activities.

Not all remaining items were strong-performing, and several items were retained that were very easy or very hard or that had a low corrected point-biserial. Most were moderate to strong items, however. The final alpha for the set of 101 scored items was .74. The difficulties ranged from 31.1% to 96.6%, with a mean of 71.6%. The r_{pb}^* values ranged from $-.04$ to $.41$, with a mean of $.15$. Note that with a small sample of candidates, negative point-biserial values are not necessarily a sign of a problematic item.

Table 11 shows the scored CHRP ELE's final fit to the domain weighting. Table 12 shows the same for cognitive level, and Table 13 shows the same for item type. The exam fit on all dimensions.

The group endorsed the final set of items for use in scoring the January 2019 CHRP ELE candidates.

Table 11: Domain fit for final scored items

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	43	46% ± 5%	42–51	—
B Employer Obligations	36	33% ± 4%	30–37	—
C Regulations and Legislation	22	21% ± 3%	19–24	—
TOTAL	101		101	—

Table 12: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	10	10% ± 3%	8–13	—
Application	62	60% ± 10%	51–70	—
Critical thinking	29	30% ± 10%	21–40	—
TOTAL	101		101	—

Table 13: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	27	25% ± 3%	23–28	—
Case	74	75% ± 3%	73–78	—
TOTAL	101		101	—

Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014)⁵ and Livingston and Kim (2009),⁶ was used to establish the pass mark for the January 2019 CHRP ELE. The goal of this process was to set a

⁵ Kolen, M.J., & Brennan, R.L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer.

⁶ Livingston, S.A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, 46, 330-343.

pass mark for the January 2019 CHRP ELE that would be equivalent to that set for previous administrations; that is, to 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 ELE was last set after the January 2018 offering of the CHRP ELE using the Modified Angoff and Bookmark methods. Specific information on the standard-setting session is provided in the Technical Report issued for the January 2018 administration.

Three equating procedures were conducted back to different administrations (January, May and September 2018). The intention following these 3 equating runs was to average them to arrive at a final pass mark for the January 2019 CHRP ELE. These administrations were chosen because they were the only available administrations under the new standard set for January 2018. The equating run to the May 2018 administration was unplanned, but conducted as a result of inconsistent results shown in the equating to the September 2018 administration.

Equating Back to the September 2018 Administration

Linear equating (Tucker) was the chosen method for setting the pass mark and it was conducted once key validation was complete. Linear equating is the primary method considered with more than 100 candidates; equipercentile equating would have been considered with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc 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 that adhered to the blueprint. 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 0.72 and a mean corrected point-biserial of .17.

Table 14 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are reasonably aligned with the targets and reflect the scope and approximate weighting across the full exam.

Table 14: Anchor item fit to blueprint – To September 2018

	Area	Actual	Target
A	Employment Contracts and Terminations	43%	46%
B	Employer Obligations	37%	33%
C	Regulations and Legislation	20%	21%

The mean, Tucker, Levine observed-score, circle arc, and Braun-Holland methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker was considered the primary method. By one of the selection metrics Levine would have been chosen, and the sample sizes are close to the 100 threshold which would lead to consideration of mean and circle arc methods. As such, though Tucker was favoured, the analysis continued with attention to prevailing patterns of results.

Table 15 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 population taking the January 2019 CHRP ELE scored higher than the population taking the September 2018 CHRP ELE (71.8% vs. 68.7%, respectively; $t(222)=1.99, p<.05$). Because the January 2019 CHRP ELE candidates were of greater ability (based on the anchors), they should have a higher pass rate (though with small samples the tails of the distribution will not be regular).

The equating analysis generally bears this out (Table 16), but with a wider than usual range of potential pass marks. The methods show a pass mark of 52 to 55. Given the sample sizes involved, Tucker or Levine observed would be the primary methods under consideration, though the testing volumes are close to the threshold for prioritizing the low volume methods (i.e., mean and circle arc). The Tucker method showed the lowest estimated equating error (between Tucker and Levine observed) and would be chosen based on general rules of thumb for deciding between the Tucker and Levine observed methods.

Table 15: Equating parameter table – To September 2018

		Sep. 2018	Jan 2019
	n	105	119
	Scored items	98	101
Mean score	Total	71.7%	71.6%
	Anchors	68.7%	71.8%

Table 16: Equating outcome table – To September 2018

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equated Sep. 2018	54.72	55	91.4%	95.2%
Tucker	53.66	54	98.5%	98.3%
Levine observed	51.43	52	99.2%	99.2%
Mean	54.93	55	98.5%	98.3%
Circle Arc 1	54.29	55	98.5%	98.3%
Circle Arc 2	54.21	55	98.5%	98.3%
Braun-Holland	54.40	55	98.5%	98.3%

Equating Back to the January 2018 Administration

Linear equating (Tucker) was the chosen method for setting the pass mark and it was conducted once key validation was complete. Linear equating is the primary method considered with more than 100 candidates; equipercentile equating would have been considered with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc 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 that adhered to the blueprint. 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 0.72 and a mean corrected point-biserial of .18.

Table 17 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are aligned with the targets and reflect the scope and weighting across the full exam.

Table 17: Anchor item fit to blueprint – To January 2018

	Area	Actual	Target
A	Employment Contracts and Terminations	45%	46%
B	Employer Obligations	33%	33%
C	Regulations and Legislation	21%	21%

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

Table 18 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 population taking the January 2019 CHRP ELE scored marginally lower than the population taking the January 2018 CHRP ELE (71.9% vs. 72.7%, respectively; $t(228)=0.57$, *ns*). Because the January 2019 CHRP ELE candidates were of marginally lower ability (based on the anchors, non-significance notwithstanding), they should have a slightly lower pass rate (though, again, the tails of the distribution will be more erratic with small samples).

The equating analysis approximates this, though the pass rate is up marginally because of a lack of very low scorers (Table 19). All methods show a pass mark of 57 to 59. Given the sample sizes involved, Tucker or Levine observed would be the primary methods under consideration. The Tucker method showed the lowest estimated equating error and would be chosen based on general rules of thumb for deciding between the Tucker and Levine observed methods.

Table 18: Equating parameter table – To January 2018

		Jan. 2018	Jan. 2019
	n	111	119
	Scored items	100	101
Mean score	Total	73.0%	71.6%
	Anchors	72.7%	71.9%

Table 19: Equating outcome table – To January 2018

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Ang/Book Jan 2018	56.82	57	95.8%	97.3%
Tucker	57.62	58	96.2%	97.5%
Levine observed	58.37	59	95.5%	96.6%
Mean	56.59	57	97.0%	98.3%
Circle arc 1	56.73	57	97.0%	98.3%
Circle arc 2	56.72	57	97.0%	98.3%
Braun-Holland	57.77	58	96.2%	97.5%
Braun-Holland	54.40	55	98.5%	98.3%

Equating Back to the May 2018 Administration

Linear equating (Tucker) was the chosen method for setting the pass mark and it was conducted once key validation was complete. Linear equating is the primary method considered with more than 100 candidates; equipercenile equating would have been considered with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc 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 that adhered to the blueprint. 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 0.75 and a mean corrected point-biserial of .17.

Table 20 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are aligned with the targets and reflect the scope and weighting across the full exam.

Table 20: Anchor item fit to blueprint – To May 2018

	Area	Actual	Target
A	Employment Contracts and Terminations	48%	46%
B	Employer Obligations	32%	33%
C	Regulations and Legislation	20%	21%

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

Table 21 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 population taking the January 2019 CHRP ELE scored marginally lower than the population taking the May 2018 CHRP ELE (75.2% vs. 75.6%, respectively; $t(230)=0.25$, *ns*). Because the January 2019 CHRP ELE candidates were of marginally lower ability (based on the anchors, non-significance notwithstanding), they should have a slightly lower pass rate (though, again, the tails of the distribution will be more erratic with small samples and the difference was less than half a percentage point).

The equating analysis approximates this, though the pass rate is up marginally because of a lack of very low scorers (Table 22). All methods show a pass mark of 57. Given the sample sizes involved, Tucker or Levine observed would be the primary methods under consideration. The Tucker method showed the lowest estimated equating error and would be chosen based on general rules of thumb for deciding between the Tucker and Levine observed methods.

Table 21: Equating parameter table – To May 2018

		May 2018	Jan. 2019
N		113	119
Scored items		100	101
Mean score	Total	71.3%	71.6%
	Anchors	75.6%	75.2%

Table 22: Equating outcome table – To May 2018

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equating May 2018	54.85	55	97.5%	97.3%
Tucker	56.09	57	97.0%	98.3%
Levine observed	56.94	57	97.0%	98.3%
Mean	56.04	57	97.0%	98.3%
Circle arc 1	56.07	57	97.0%	98.3%
Circle arc 2	56.07	57	97.0%	98.3%
Braun-Holland	56.04	57	97.0%	98.3%

Combined Results

Table 23 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. However, the results of September are clearly lower to a substantial degree than the January and May 2018 results. Further, the score distribution for the September administration was oddly shaped and seeming out of keeping with a typical administration. As such, the most prudent pass mark was identified by the simple arithmetical mean (56.857) of the January and May 2018 equating runs.

Using the established convention for this testing program, the combined pass mark would be rounded up to a cut score of 55. The resulting pass rate for first-time candidates is slightly higher than that observed for the May and January 2018 CHRP ELE: 97.3% for both. The pass rate for all candidates was 97.0% which was just slightly lower than the 97.5% in May 2018 and slightly higher than the 95.8% in January 2018. Given that the differences between the samples (based on the anchors) were marginal and nonsignificant, little change was expected in pass rates. See Table 24 for historical pass rates.

The final pass mark value, and the process used to derive it, was presented to the CHRP EVC (Table 25) via teleconference on January 28, 2019. No concerns were raised regarding the pass mark or pass rate. The panel formally approved the pass mark (which was presented along with the consequent pass rate data) for recommendation to HRP. The HRP Registrar accepted the recommended pass mark and so the pass mark was formally established.

Table 23: Equating outcome table – Combined results

	Jan. 18	May 18	Sep. 18
Tucker	57.6	56.1	53.7
Levine observed	58.4	56.9	51.4
Mean	56.6	56.0	54.9
Circle arc 1	56.7	56.1	54.3
Circle arc 2	56.7	56.1	54.2
Braun-Holland	57.8	56.0	54.4

Table 24: Historical pass rates

	Pass rate	
	All	First-time
Jan. 17	94.6%	95.7%
May 17	94.7%	95.2%
Sep. 17	95.6%	95.9%
Jan. 18	95.8%	97.3%
May 18	97.5%	97.3%
Sep. 18	91.4%	95.2%
Jan. 19	97.0%	98.3%

Table 25: CHRP Examination Validation Committee – Pass mark approval

Member	Credential	Years of Relevant Experience	Years on EVC	Industry
✓ Sunday Ajao	CHRL	15–20	2	Banking/Finance
✓ Roxanne Chartrand	CHRL	20–29	1	Insurance
Claire Chester	CHRL	10–14	2	Regulation/CPA
Tanya Gopaul	CHRL	10–15	2	Banking
Jean Lazarus	CHRL	15–19	2	Health services
Suman Seth	CHRL	15–19	1	Public sector
Kriss Stone	CHRP	10–15	2	Real estate
✓ Ielean Tait	CHRL	15–20	2	Environmental
Karen Weiler	CHRL	20–29	2	Software/ Communications
Alyssa Young	CHRL	5–9	2	Non-profit

✓ Participated in the session.

Scoring

To finalize the scoring, candidates who were not included in the item and form analyses were reinserted into the dataset. Scores for each of the 3 domain areas were also computed for each candidate. An Excel file with the final candidate results was provided to HRP.

Table 26 provides the means and standard deviations for the domains and for the total score, using all candidates who took the January 2019 CHRP ELE. Table 27 provides the correlations between each domain. Caution should be exercised in interpreting differences between correlations. Variation can be explained largely by the number of items making up each domain score. That is, domain C has fewer items and shows lower correlations with the other domains. Figure 3 shows the distribution of scores for all candidates, along with the pass mark.

Table 26: Total and domain scores for all candidates

Domain	Percentage	Mean	SD*
A Employment Contracts and Terminations	71%	30.6	4.0
B Employer Obligations	67%	24.2	3.7
C Regulations and Legislation	76%	16.8	2.2
Total score	70.9%	71.6	8.1

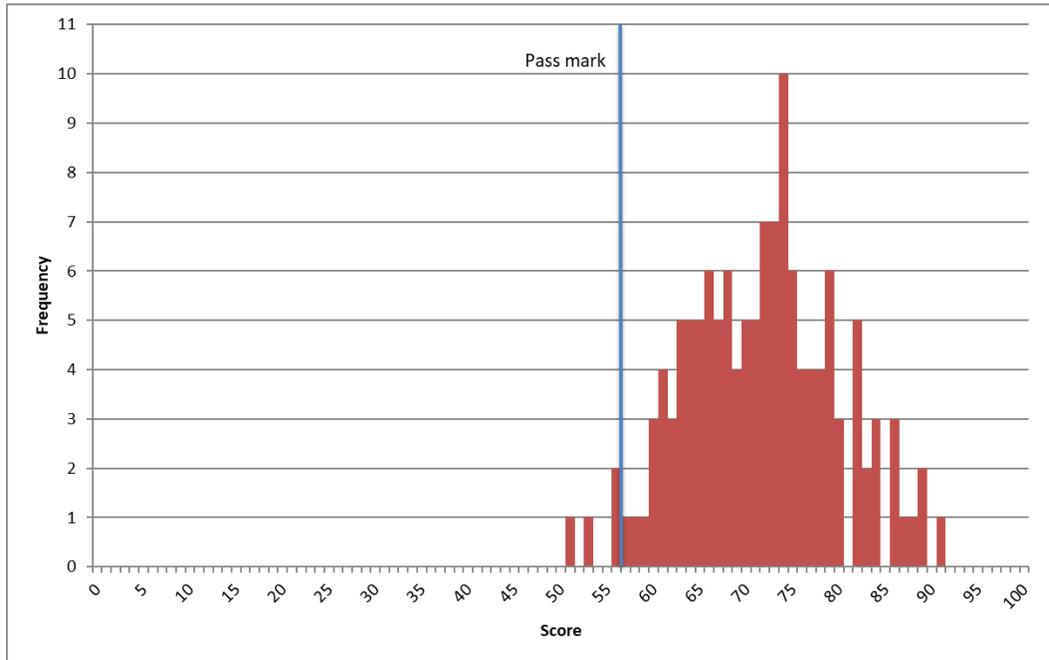
*SD = Standard deviation.

Table 27: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.53	.44
B			.44
C			

*See Table 26 for the full name of each functional area.

Figure 3: Score distribution for all candidates



Key Examination Metrics

Table 28 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.

Note that as of January 2018 the time limit for candidates was increased from 3 hours to 3½ hours.

Table 28: Key examination metrics – Candidates included in analysis only

Index	January 2019	September 2018	May 2018	January 2018	September 2017
Scored items	101	98	100	100	100
Candidates	119	105	113	111	123
Mean	72.29 (71.6%)	70.30 (71.7%)	71.30 (71.3%)	73.03 (73.0%)	71.54 (71.5%)
Median	73 (72.3%)	72 (73.5%)	71 (71.0%)	73 (73.0%)	72 (72.0%)
Skewness	0.005	-0.593	-0.350	-0.432	-0.530
Kurtosis	-0.244	-0.121	0.281	-0.023	0.073
Range	51–91 (50.5– 90.1%)	49–84 (50.0– 85.7%)	50–88 (50.0– 88.0%)	49–90 (49.0– 90.0%)	44–90 (44.0– 90.0%)
Standard deviation	8.01	8.04	7.80	8.22	9.33
Cronbach's alpha	.74	.75	.72	.76	.81
Mean r_{pb}^*	.15	.15	.14	.15	.18
SEM ⁱ	4.11	4.06	4.09	4.04	4.08
SEM at the pass mark	4.59	4.55	4.57	4.59	4.61
Decision consistency (uncorrected) ⁱⁱ	.95	.94	.95	.95	.96
Perceived fairness ⁱⁱⁱ	63%	64%	47%	59%	54%
Pass mark	56.857	54.718	54.846	56.817	53.363
Effective pass mark	57	55	55	57	54
Pass rate	98.3%	95.2%	97.3%	97.3%	95.9%

ⁱSEM = standard error of measurement.

ⁱⁱSubkoviak method.

ⁱⁱⁱBased on responses to the post-examination survey for all candidates.

Related Development Activities

Since the last administration of the CHRP ELE in September 2018, the following exam development activities have taken place.

Item Validation

To provide sufficient items for upcoming administrations, a validation session was held November 19–20, 2018, at HRPAs offices. The CHRP EVC members who participated are listed in Table 29. This session involved the review of CKE 1 items as well.

Table 29: CHRP Examination Validation Committee – Item validation

Member	Credential	Years of Relevant Experience	Years on EVC	Industry
Sunday Ajao	CHRL	15–20	2	Banking/finance
✓ Claire Chester	CHRL	10–14	2	Regulation/CPA
✓ Tanya Gopaul	CHRL	10–15	2	Banking
✓ Jean Lazarus	CHRL	15–19	2	Health services
Kriss Stone	CHRP	10–15	2	Real estate
Ielean Tait	CHRL	15–20	2	Environmental
✓ Karen Weiler	CHRL	20–29	2	Software/ communications
✓ Alyssa Young	CHRL	5–9	2	Non-profit

✓ Participated in the session.

Two new members were added to the committee subsequent to this session.

The EVC members received advance materials outlining the following:

- Purpose of the session
- Description of the CHRP credential
- CKE 1 and CHRP ELE blueprints
- Criteria for good test items
- Validation process
- Relevant legislation

At the session, abbreviated training was provided as these committee members were already informed about the credentials and general process from previous sessions. This training included a primer on psychometrics to orient the committee to the item statistics they would be working with.

For each item, the committee was asked to either

- Validate the item for use in the next 2 years to make decisions about who would be certified as an HR professional in Ontario (at the CHRP level),
- Move the item to the CHRL ELE bank,
- Revise the item to make it suitable for use, or
- Declare the item unsound and send it back for revision or removal from the bank.

The bulk of the session saw the committee members reviewing items independently and submitting their assessments in blocks of approximately 15–25 items. Those assessments were tabulated and any items that were not validated as is by the full committee were discussed until there was agreement on changes and the future use of the item. Items with problematic statistics were reviewed and edited as a group.

The panel members reviewed and validated 100 items (including 14 cases) as suitable for the CHRP ELE, designated 0 items as suitable only for the CHRL ELE, and rejected 1 item as unfixable. About 20% of items were edited before validation, with most of these edited items being those with a history of problematic statistics.

Item Writing

To fill gaps in the bank and renew content, item writing was conducted in October–December 2018. Item writers (see Table 30) were identified by HRP A and trained in a remote session by Wickett on October 29, 2018. Items written in this process were considered eligible for both the CHRP ELE and the CHRL ELE.

Table 30: Item writers

Writer	Credentials	Years of Relevant Experience	Industry
Susan Haywood	MA, CHRL, SCP	15 years in HR, 3 years teaching HR	Consultant, Education, Professor - Algonquin College
Gail Lawrence	CHRL, MSc.Mgt	15 years teaching in HR, 25 years HR consulting	Profit and Not for Profit
Paula Matkowski	CHRP, CHRL	21 years in HR, 9 years teaching HR	Manufacturing, Public Sector, Family-Owned, Union, Non-Union, Education – Part Time Professor at Niagara College
Alanna Twohey	B.A., LL.B	6 years teaching in HR, 10 years practicing in HR Law	Legal Services, Education – Professor at Algonquin College

The item writers were provided with training via teleconference, and received printable files covering the main elements of the training. The general guidance for writing quality multiple choice items was drawn primarily from Haladyna & Rodriguez (2013).⁷

Each item writer was selected based on expertise in identified topics from the ELE blueprints, and they were assigned items within those topics as per their stated preferences. Item writers were assigned 15–30 items each to write, for a total of 90 items (one writer did not fully complete their items, and so 83 new items were obtained).

The item writers had access to the style guide that governs language usage on the HRP A exams and were provided with recent text books and employment law 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 review, validation and upon use with candidates. This generally required several iterations per item.

Once all items were drafted and declared complete, they were sent a certified professional editor for editorial review. Items were adjusted based on this input and comments noted if future reviewers would need to attend to specific content concerns. Further, HRP A's reference librarians verified the authoritative source for each item and any requests for alternate sources were sent back to the item writer to produce.

Item Review

Following the item writing exercise in October–December 2018 there was need for group review of those items before moving them to formal validation and use on the CHRP ELE. The group had 247 items plus 15 case texts for consideration (taken from newly written items, supplemented with other unreviewed items in the bank and items requiring revision required to fill gaps in the bank). The group was not expected to be able to review all items in the time available. Items reviewed in this process were considered eligible for both the CHRP ELE and the CHRL ELE unless otherwise indicated by the group.

The 2½-day review session was held January 8–10, 2019 at HRP A offices. The panel members who participated are shown in Table 31.

⁷ Haladyna, T. M., & Rodriguez, M.C. (2013). *Developing and validating test items*. New York, NY: Routledge.

Table 31: Panel for item review session

Reviewer	Credentials	Years of Relevant Experience	Industry
Joel Andre	CHRL, M.HRM	11	Transportation
Janet Brooks	CHRL, CHRE, MBA	38	Consulting
Tanya Dacres	CHRL	16	Marketing/Technology
Bonnie Hall	CHRL, CPM	16	Food manufacturing
Arvinder Hanspal	CHRL	15	Manufacturing
Jasmyn Kozlowsky	CHRL	12	Manufacturing
Sandra MacLeod	CHRL	15	Professional services
Heather McCarthy	CHRL	12	Not for profit
Shanta Singh	CHRL	12	Financial

The panel members received training on the review activity and context on the intent of the CHRP ELE and CHRL ELE. For the new items, the members worked primarily individually reviewing items to make sure they reflected current practice. Where panel members proposed changes, these were discussed by the group before implementation. For the items with statistical history or comments from past panels on quality, those items were reviewed as a group with input from Wickett on the nature of the issue with the item and guidance on what type of revision may be helpful to improving the effectiveness of the item. The group formulated and approved changes to these items.

The panel members reviewed and approved 133 items as suitable for the CHRP ELE and CHRL ELE, moved 2 items to the CKE 2 bank, and rejected 14 items. Of the approved ELE items, 62 saw text changes to the stem and/or options before approval.

The group reported confidence in the final assessments made by the group and satisfaction overall with the session.

The items were updated in the bank, and those that were approved were deemed ready for validation before use on future examinations.

Appendix

Blueprint

CHRP Employment Law Examination Blueprint

Human Resources Professionals Association

Version 2.1

Approved by CHRP Exam Validation Committee April 10, 2018

Approved by HRPA Registrar April 11, 2018

Effective September 2018 administration

Credential

Passing the CHRP Employment Law Examination is a requirement for certification of CHRP candidates.

Purpose

The CHRP ELE assesses whether a candidate has the ability to make effective decisions when presented with HR situations where comprehension of laws and regulations is centrally relevant, at the CHRP level, in Ontario.

Structure

The structural variables provide high level guidance as to what the examination will look like. These appear in Table 32.

Table 32: CHRP Employment Law Examination Blueprint Structural Variables

Item types	75% Case-based 3-option multiple choice (15-20 single scenarios tied to 4-6 test items each)
	25% Independent 3-option multiple choice
Length	110 total items
	8–12 experimental items
Duration	Up to 3½ hours
Delivery mode	Computer based testing in proctored test centres
Frequency	3 windows per year

Content Weighting

The topic weights were set through a survey of employment lawyers on the most typical situations where employment-related issues are escalated to legal proceedings.

Categories are:

- A. Employment Contracts
- B. Employer Obligations
- C. Regulations and Legislation

Within each Category, the Topics are:

- A. Employment Contracts
 - A1 Termination
 - A2 Contracts
 - A3 Employee Benefits and Perquisites
- B. Employer Obligations
 - B1 Duty to Accommodate
 - B2 Misconduct in the Workplace
 - B3 Common Law
 - B4 Sale of Business
- C. Regulations and Legislation
 - C1 Employment Standards Act
 - C2 Occupational Health and Safety Act
 - C3 Jurisdiction
 - C4 Pay Equity Act
 - C5 Canada Labour Code

The full blueprinted list of Categories, Topics and Subtopics, along with their weighting, appears in Table 33.

Table 33: CHRP Employment Law Examination Blueprint Content Weights

Category Weight	Topic Weight	Topic	Subtopic Weight
46%	A. Employment Contracts and Terminations		
	28%	A1. Termination	
		A1.1 Termination with or without cause	8%
		A1.2 Termination pay, termination notice, and pay in lieu of notice	6%
		A1.3 Continuation of benefits to employee after termination	5%
		A1.4 Severance pay entitlements	5%
		A1.5 What type of income is considered part of terminated employee's salary	2%
		A1.6 Whether or not it is legal to lay off an employee	1%
	A1.7 When and how to lay off an employee	1%	
	11%	A2. Contracts	
		A2.1 Contracts and employment agreements	9%
	A2.2 Collective bargaining contracts	2%	
	7%	A3. Employee Benefits and Perquisites	
		A3.1 Vacation time, vacation pay and bonuses	5%
A3.2 Overtime exemptions		2%	
33%	B. Employer Obligations		
	16%	B1. Duty to Accommodate	
		B1.1 Mental health or physical disabilities	9%
		B1.2 Discriminatory grounds (such as family status, age, marital status, etc.)	5%
	B1.3 The duty to accommodate until undue hardship (the threshold)	2%	
	9%	B2. Misconduct in the Workplace	
		B2.1 Dealing with harassment and violence in the workplace	5%
		B2.2 HR professional approach to dealing with discipline	2%
	B2.3 Workplace investigations	2%	
	6%	B3. Common Law	
		B3.1 Including consideration of Common Law principles	5%
	B3.2 Employers' obligations under Common Law	1%	
	2%	B4. Sale of Business	
B4.1 The effects of the sale of the business		2%	
21%	C. Regulations and Legislation		
	10%	C1. Employment Standards Act	
		C1.1 How to properly interpret the <i>Employment Standards Act, 2000</i>	5%
C1.2 Probation period under <i>Employment Standards Act, 2000</i>	2%		

	C1.3 Different leaves permitted under the <i>Employment Standards Act, 2000</i>	2%
	C1.4 Employers' obligations under <i>Employment Standards Act, 2000</i>	1%
4%	C2. Occupational Health and Safety Act	
	C2.1 Making policies that are compliant with the <i>Occupational Health and Safety Act, 1990</i>	2%
	C2.2 Ministry of Labour's rights under the <i>Occupational Health and Safety Act, 1990</i>	2%
4%	C3. Jurisdiction	
	C3.1 The difference between federal and provincial legislations	2%
	C3.2 Determining governing legislation when the organization is interprovincial	2%
2%	C4. Pay Equity Act	
	C4.1 Application of <i>Pay Equity Act, 1990</i>	2%
1%	C5. Canada Labour Code	
	C5.1 Employers' obligations under <i>Canada Labour Code, 1985</i>	1%

Note: Reasonable ranges around the Topic weights are employed.

Cognitive Level

The cognitive level weights are based on Bloom's taxonomy. The purpose of this weighting is generally to ensure that an examination does not unintentionally over-focus on specific types of items, and to provide candidates with a range of items (in approximate proportion) that reflects the cognitive operations they must apply on the job. The weights appear in Table 34.

Table 34: CHRP Employment Law Examination Blueprint Cognitive Level Weights

Level	Weight	Range
Knowledge	10%	+/- 3%
Application	60%	+/- 10%
Critical Thinking	30%	+/- 10%

Miscellaneous Guidance

Guidance is not considered binding on the examination, but is used in item development and form development to help create balanced forms.

- Where scenarios or test items include a workplace, the workplace allocation will be as follows:
 - For profit enterprise, 60% (+/- 10%)
 - Government, 20% (+/- 5%)
 - Not-for-profit, 20% (+/- 5%)
- 20% (+/- 10%) of workplaces mentioned in scenarios and test items will be unionized.

3. 10% (+/- 5%) of employers mentioned in scenarios and test items will have physical locations in more than one Canadian province.
4. 10% (+/- 5%) of employers mentioned in scenarios and test items will have physical locations both inside and outside of Canada.