PUBLIC RELEASE

Technical Report: August 2020 CHRP ELE

Human Resources Professionals Association

19 October 2020

ICKETT
MEASUREMENT SYSTEMS

Contents

Executive Summary **5**
Administration **6**
Form Setting 6
Testing Window 7
Analysis **9**
Data Cleaning and Integrity Checks 9
Post-Examination Survey 11
Initial Analysis 14
Key Validation 15
Establishing the Pass Mark: Equating 18
Scoring 26
Key Examination Metrics **28**
Related Development Activities **29**
Validation 29
Appendix **31**
Blueprint 31

List of Tables

Table 1: Domain fit at administration 6
Table 2: Cognitive level fit at administration 7
Table 3: Content-related post-examination survey questions* 12
Table 4: General administration-related post-examination survey questions* 12
Table 5: Testing location 13
Table 6: Preferred location (live remove proctoring candidates) 13
Table 7: Reason for choosing own location (live remove proctoring candidates) 13
Table 8: Evaluation of testing experience (live remove proctoring candidates) 13
Table 9: Value in future candidates being able to test from their own location (live remote proctoring candidates) 14
Table 10: Able to write at a convenient location (test centre candidates) 14
Table 11: Value in future candidates being able to test from their own location (test centre candidates) 14
Table 12: Initial examination statistics 15
Table 13: CHRP Examination Validation Committee – Key validation 16
Table 14: Domain fit for final scored items 17
Table 15: Cognitive level fit for final scored items 17
Table 16: Item type fit for final scored items 17
Table 17: Anchor item fit to blueprint – To January 2020 19
Table 18: Equating parameter table – To January 2020 19
Table 19: Equating outcome table – To January 2020 20
Table 20: Anchor item fit to blueprint – To May 2019 21
Table 21: Equating parameter table – To May 2019 21
Table 22: Equating outcome table – To May 2019 22
Table 23: Anchor item fit to blueprint – To May 2019 23
Table 24: Equating parameter table – To September 2019 23
Table 25: Equating outcome table – To September 2019 24
Table 26: Equating outcome table – Combined results 25
Table 27: Historical pass rates 25
Table 28: CHRP Examination Validation Committee – Pass mark approval 26

Table 29: Total and domain scores for all candidates....26
Table 30: Correlations between functional area scores for all candidates 27
Table 31: Key examination metrics – Candidates included in analysis only....28
Table 32: CHRP Examination Validation Committee members – Validation....29
Table 33: CHRP Employment Law Examination Blueprint Structural Variables....31
Table 34: CHRP Employment Law Examination Blueprint Content Weights 33
Table 35: CHRP Employment Law Examination Blueprint Cognitive Level Weights....34

List of Figures

Figure 1: Examination time distribution for all candidates....10
Figure 2: Candidate volume and score trends across testing window 10
Figure 3: Score distribution for all candidates....27

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).

Special note: This administration was originally scheduled to occur in May 2020 but was moved to August 2020 because of COVID-19 closures.

The CHRP Employment Law Exam (CHRP ELE) was administered to 126 candidates using computer-based testing via live remote proctoring and at Prometric test centres August 10–16, 2020, 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[2] (*n*=116) was 71.5 (71.5%), and for all candidates it was 70.6 (70.6%), out of 100 validated items for scoring. Reliability was marginal at .78 (noting that there is range restriction with these candidates; disattenuated reliability is estimated at .83). The final set of scored items adhered to the blueprint parameters.

The pass mark was set using equating back to the May 2019, September 2019 and January 2020 administrations, yielding an integer pass mark of 54. 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 96.6% and a pass rate for all candidates of 96.0%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards[3] 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 HRPA 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 Suman Seth and Alyssa Young (CHRP Examination Validation Committee members) in a remote session held February 21, 2020.

The final form composition for the August 2020 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, HRPA makes use of previously validated and administered test forms along with new test forms, in addition to employing other mechanisms to maintain the integrity of the exams and candidate scores.

Table 1: Domain fit at administration

	Domain	Actual Items	Target Range	Target Items	Variance
A	Employment Contracts and Terminations	50	46% ± 5%	46–56	—
B	Employer Obligations	38	33% ± 4%	32–40	—
C	Regulations and Legislation	22	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	62	60% ± 10%	55–77	—
Critical thinking	38	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 using live remote proctoring and at Prometric test sites primarily in Ontario. The testing window was August 10–16, 2020, inclusive, and 126 candidates wrote the exam.[5]

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[6]) were employed for both methods. Candidates were allowed up to one unscheduled 10-minute break during the examination (the examination timer was not stopped during this break).

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*

[5] Due to technical difficulties requiring the rescheduling of some candidates, testing continued on August 17.

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

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 February 14, 2020.

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, 50 minutes (down 17 minutes from January 2020). Note that minor variances in how time is recorded have been seen over the past few administrations, and so it is unclear if time differences are reflecting a technical issue or an actual change in candidate behaviour relative to the examination.

Fourteen candidates (11%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 4 candidates (3%) 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. Compared with the January 2020 administration, there was a decrease in these values. Note that because they have access to legislation, candidates may take more time than intended by researching more answers. This may generally skew time metrics higher.

The correlation between scores on the 110 items and time spent writing the examination was negligible at a value of −.05, suggesting that time constraints were not generally related to candidate performance.

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 2 days was a decrease of 1.2 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. As this was a shorter administration window that also included live remote proctoring, the normal pattern of increasing volume at the end of the window was not expected nor observed.

Figure 1: Examination time distribution for all candidates

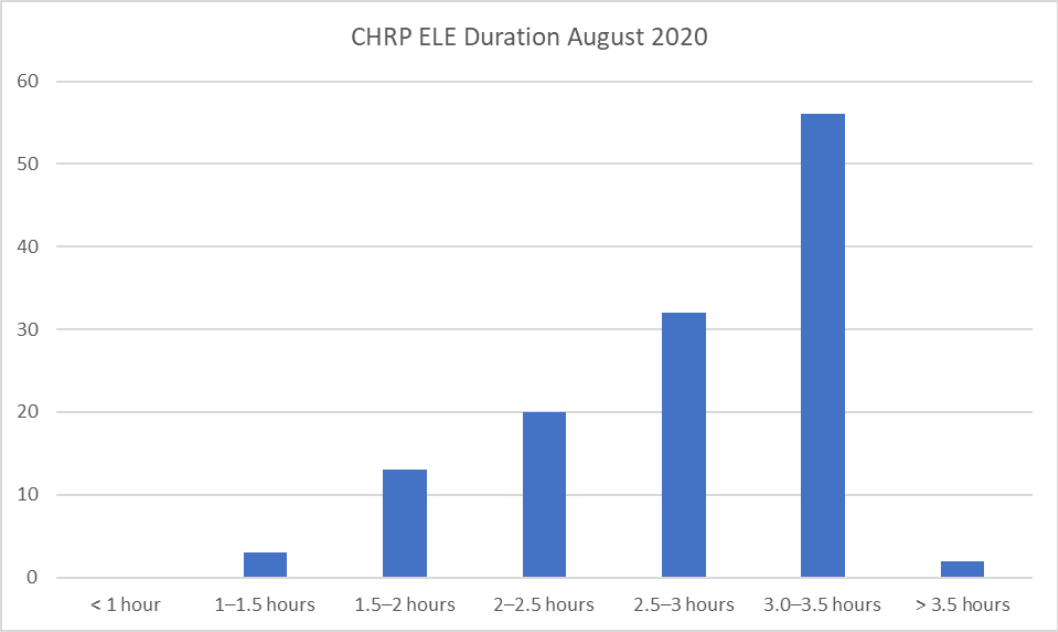
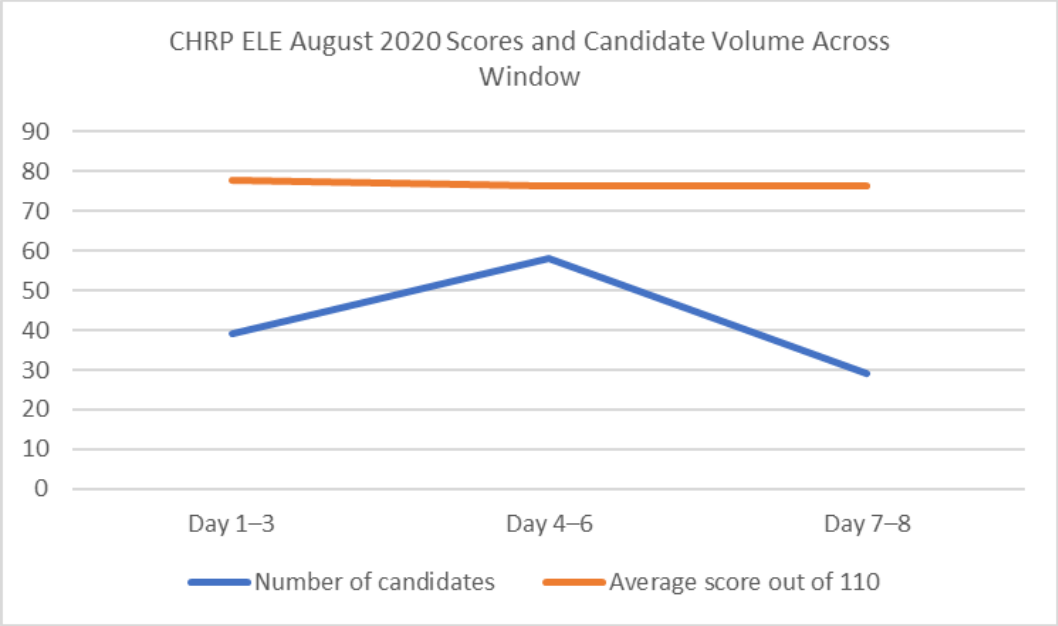


Figure 2: Candidate volume and score trends across testing window



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. One candidate was 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, 4 candidates were removed from analyses.

Candidates who had failed a previous employment law examination (CHRP ELE or CHRL ELE; n=7) scored lower than did those who had not (66.4 and 77.4, respectively, on the full exam of 110 items; $t(8)=4.63$, $p<.01$). In keeping with standard procedures, these candidates were removed from subsequent analyses. The CHRP ELE analysis proceeded with 116 candidates (note that one of the repeat candidates was also flagged in the analysis outlined in the previous paragraph).

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; 124 candidates responded (response rate, 98%). The survey was revised for the August administration to collect information specifically applicable to the use of live remote proctoring.

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

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

Question		SA	A	N	D	SD	Score	Agreement
1.	The time allotted for this examination was sufficient.	56	39	5	16	8	3.96	**77%**
2.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	33	66	13	12	0	3.97	**80%**
3.	I feel I was adequately prepared to write this examination.	14	64	27	19	0	3.59	**63%**
4.	The questions in the examination were clearly written.	11	62	32	16	3	3.50	**59%**
5.	The terminology used in the examination was accurate.	15	87	15	6	0	3.90	**83%**
6.	The situations presented in the examination were realistic.	25	85	11	3	0	4.06	**89%**
7.	The questions in the examination reflected the Employment Law Examination blueprint.	12	66	31	12	3	3.58	**63%**
8.	The examination was a fair assessment of my ability.	9	61	30	19	5	3.40	**56%**

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

Table 4: General administration-related post-examination survey questions*

Question		SA	A	N	D	SD	Score	Agreement
9.	I was able to book to write the examination at a time that was convenient for me.	35	62	8	10	9	3.84	**78%**
10.	I was well informed about the examination rules and regulations.	51	65	7	1	0	4.34	**94%**
11.	Proctors enforced the exam-day rules.	70	49	4	1	0	4.52	**96%**
12.	Proctors were professional and courteous.	78	40	5	1	0	4.57	**95%**
13.	The tutorial helped me understand how to complete the examination on the computer.	57	55	9	2	0	4.36	**91%**
14.	The legislation and case texts were easy to access during the examination.	37	50	15	16	6	3.77	**70%**
15.	Navigation through the examination was easy and intuitive.	45	69	10	0	0	4.28	**92%**

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

Candidates were asked where they wrote the examination, and based on their response the questions that followed differed (see Table 5).

Table 5: Testing location

Response	Count	%
Test centre	31	25%
Own location	93	75%

Candidates who indicated they tested in the own location (via live remote proctoring) responded to questions shown in Table 6 through Table 9. These candidates were largely split on whether they preferred using their own location versus a test centre, though they 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 were a motivating factor for many in choosing live remote proctoring (shown primarily in the 'other' write in responses), and lack of open test centres was also a main driver.

Table 6: Preferred location (live remove proctoring candidates)

Response	Count	%
I preferred using my own location.	36	39%
I preferred going to a test centre.	44	47%
I have no preference.	13	14%

Table 7: Reason for choosing own location (live remove proctoring candidates)

Response	Count	%
No test centres were open in my area.	29	31%
I preferred to avoid being around other people.	20	22%
I liked the convenience of not having to travel to a test centre.	23	25%
I felt like I would perform better in my own environment.	6	6%
Other (please specify)	15	16%

Table 8: Evaluation of testing experience (live remove proctoring candidates)

	Count	%
Very positive	15	16%
Positive	41	44%
Neutral	17	18%
Negative	14	15%
Very negative	6	6%

Table 9: Value in future candidates being able to test from their own location (live remote proctoring candidates)

Response	Count	%
Yes	85	92%
No	7	8%

Candidates who indicated they tested in a test centre responded as shown in Table 10 and Table 11. These candidates indicated some challenges in being able to write in their preferred location, though given the constraints related to COVID-19 closures, they were arguably more positive than would have been expected. These candidates were also generally supportive of HRPA continuing to offer the option of writing using live remote proctoring in the future.

Table 10: Able to write at a convenient location (test centre candidates)

	Count	%
Strongly agree	6	20%
Agree	14	47%
Neither agree nor disagree	0	0%
Disagree	5	17%
Strongly disagree	5	17%

Table 11: Value in future candidates being able to test from their own location (test centre candidates)

Response	Count	%
Yes	23	77%
No	7	23%

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.

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 12.

Table 12: Initial examination statistics

Index	CHRP ELE
Items	110
Total candidates	126
Candidates in analysis	116
Mean	77.8 (71%)
Standard deviation	9.1
Range	54–100 (49–91%)
Cronbach's alpha	.78
Disattenuated alpha	.85
Mean r_{pb}*	.15

A simple comparison between scores obtained by test centre candidates (mean score of 68.9%) and live remote proctoring candidates (mean score of 70.1%) was made to evaluate if there was any problematic difference in performance. The small number of candidates means this analysis is inconclusive, but a lack of significant effect ($t(124)=0.65$, *ns*) is at least supportive of there being no overall difference in candidate performance. This will be monitored across future administrations.

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

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 generally ranked from worst performing to best performing accordingly.

Key Validation

Key validation was conducted via web meeting on August 19, 2020, using members of the CHRP Examination Validation Committee (EVC). The group (Table 13) 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 13: CHRP Examination Validation Committee – Key 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	10–15	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 group was informed that test reliability, as measured by Cronbach's alpha, was .78 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 and were provided with the disattenuated value of .85 as an estimate of the true reliability of these test scores. 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. Further, historic data on items was used to help in making decisions such that items with strong statistics on past administrations were more likely to be retained for scoring. The group made decisions based on content and the data through discussion; they removed 10 items that they felt were inappropriate 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 100 scored items was .78 (disattenuated alpha was .83). The difficulties ranged from 31.9% to 95.7%, with a mean of 71.5%. The r_{pb}* values ranged from −.03 to .45, with a mean of .17. Note that with a small sample of candidates,

negative point-biserial values are not necessarily a sign of a problematic item, and items that have performed well in the past were more likely to be retained even if showing a poor point-biserial in this candidate sample.

Table 14 shows the scored CHRP ELE's final fit to the domain weighting. Table 15 shows the same for cognitive level, and Table 16 shows the same for item type. The exam fit on all dimensions.

The group approved the final set of items for use in scoring the August 2020 CHRP ELE candidates.

Table 14: Domain fit for final scored items

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	45	46% ± 5%	41–51	—
B Employer Obligations	34	33% ± 4%	29–37	—
C Regulations and Legislation	21	21% ± 3%	18–24	—
TOTAL	**100**		**100**	**—**

Table 15: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	8	10% ± 3%	7–13	—
Application	57	60% ± 10%	50–70	—
Critical thinking	35	30% ± 10%	20–40	—
TOTAL	**100**		**100**	**—**

Table 16: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	26	25% ± 3%	22–28	—
Case	74	75% ± 3%	72–78	—
TOTAL	**100**		**100**	**—**

Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014)[7] and Livingston and Kim (2009),[8] was used to establish the pass mark for the January 2020 CHRP ELE. The goal of this process was to set a pass mark for the August 2020 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 (May 2019, September 2019 and January 2020). The intention following these equating runs was to average them to arrive at a final pass mark for the August 2020 CHRP ELE. These administrations were chosen as the most recent administration (January 2020), the administration from 1 year in the past (May 2019 based on the original schedule for this exam and the premise for form setting), and the mid-point administration (September 2019) to add confidence in the final pass mark.

Equating Back to the January 2020 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 0.71 and a mean corrected point-biserial of .20.

Table 17 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.

[7] Kolen, M.J., & Brennan, R.L. (2014). *Test equating, scaling, and linking.* New York, NY: Springer.
[8] Livingston, S.A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, *46*, 330-343.

Table 17: Anchor item fit to blueprint – To January 2020

	Area	Actual	Target
A	Employment Contracts and Terminations	44%	46%
B	Employer Obligations	33%	33%
C	Regulations and Legislation	22%	21%

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 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 August 2020 CHRP ELE scored marginally higher than the population taking the January 2020 CHRP ELE (71.2% vs. 69.5%, respectively; $t(236)=1.12$, *ns*). Because the August 2020 CHRP ELE candidates were of marginally greater ability (based on the anchors, non-significance notwithstanding), they should have a marginally higher or the same pass rate (though with small samples the tails of the distribution will not be regular).

The equating analysis showed this result (Table 19). The methods all showed an integer pass mark of 53–55. Given the sample sizes involved, Tucker would be the primary method under consideration and the equated value of 53.53 was carried forward in the analysis.

Table 18: Equating parameter table – To January 2020

		Jan. 2020	Aug. 2020
	N	122	116
	Scored items	102	100
Mean score	Total	71.0%	71.5%
Mean score	Anchors	69.5%	71.2%

Table 19: Equating outcome table – To January 2020

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equated Jan. 2020	**56.13**	**57**	**92.0%**	**94.3%**
Tucker	53.53	54	*96.0%*	*96.6%*
Levine observed	53.00	53	*96.0%*	*96.6%*
Mean	54.10	55	*94.4%*	*94.8%*
Circle Arc 1	54.24	55	*94.4%*	*94.8%*
Circle Arc 2	54.24	55	*94.4%*	*94.8%*

Equating Back to the May 2019 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 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 2019

	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, and circle arc 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 August 2020 CHRP ELE scored essentially the same as the population taking the May 2019 CHRP ELE (71.6% vs. 71.8%, respectively; $t(272)=0.17$, *ns*). Because the August 2020 CHRP ELE candidates were of the same ability (based on the anchors), they should have the same pass rate (though, again, the tails of the distribution will be more erratic with small samples).

The equating analysis shows this result (Table 22). All methods show a pass mark of 52–54. 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 2019

		May 2019	Aug. 2020
	N	158	116
	Scored items	102	100
Mean score	Total	73.2%	71.5%
Mean score	Anchors	71.8%	71.6%

Table 22: Equating outcome table – To May 2019

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equated May 2019	**55.97**	**56**	**95.9%**	**96.8%**
Tucker	52.13	53	*96.0%*	*96.6%*
Levine observed	51.57	52	*96.8%*	*97.4%*
Mean	53.06	54	*96.0%*	*96.6%*
Circle Arc 1	53.50	54	*96.0%*	*96.6%*
Circle Arc 2	53.48	54	*96.0%*	*96.6%*

Equating Back to the September 2019 Administration

Because of differences in final pass mark estimates from the previous two analyses, and because of the delay of the exam to August, equating was also conducted back to the September 2019 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 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 23: Anchor item fit to blueprint – To May 2019

	Area	Actual	Target
A	Employment Contracts and Terminations	47%	46%
B	Employer Obligations	32%	33%
C	Regulations and Legislation	21%	21%

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 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 August 2020 CHRP ELE scored negligibly lower than the population taking the September 2019 CHRP ELE (71.6% vs. 72.0%, respectively; $t(212)=0.24$, *ns*). Because the August 2020 CHRP ELE candidates were of negligibly lower ability (based on the anchors, non-significance notwithstanding), they should have a negligibly smaller pass rate (though, again, the tails of the distribution will be more erratic with small samples).

The equating analysis shows this result (Table 22). All methods show a pass mark of 53–56. 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. Given that the September 2019 candidate count fell just below 100, a case could be made for selecting the mean or circle arc methods. For consistency, however, Tucker was selected.

Table 24: Equating parameter table – To September 2019

		Sep. 2019	Aug. 2020
	N	98	116
	Scored items	100	100
Mean score	Total	72.7%	71.5%
Mean score	Anchors	72.0%	71.6%

Table 25: Equating outcome table – To September 2019

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equated Sep. 2019	**56.72**	**57**	**89.5%**	**95.9%**
Tucker	54.01	55	*94.4%*	*94.8%*
Levine observed	52.99	53	*96.0%*	*96.6%*
Mean	55.82	56	*92.9%*	*94.0%*
Circle Arc 1	55.93	56	*92.9%*	*94.0%*
Circle Arc 2	55.92	56	*92.9%*	*94.0%*

Combined Results

Table 26 shows the pass mark values across the equating runs. The value highlighted in green is the one that would be selected based on sample parameters at each equating run. Though there are differences within each equating run, they converge closely on an integer value of 54. Consistent with past practices, the pass mark was identified by the weighted mean (53.108) of the Tucker values.

Using the established convention for this testing program, the averaged pass mark would be rounded up to a cut score of 54. The resulting pass rate for first-time candidates (96.6%) is the same or slightly higher than in recent administrations, which is in line with expectations from the equating runs. The pass rate for all candidates was 96.0%. See Table 27 for historical pass rates.

The final pass mark value, and the process used to derive it, was presented to the CHRP EVC (Table 28) via teleconference on August 21, 2020. 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 HRPA. The HRPA Deputy Registrar (on behalf of the Registrar) accepted the recommended pass mark and so the pass mark was formally established.

Table 26: Equating outcome table – Combined results

	May 19	Sep. 19	Jan. 20
Tucker	52.1	54.0	53.5
Levine observed	51.6	53.0	53.0
Mean	53.1	55.8	54.1
Circle arc 1	53.5	55.9	54.2
Circle arc 2	53.5	55.9	54.2

Table 27: 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%
May 19	95.9%	96.8%
Sep. 19	89.2%	95.9%
Jan. 20	92.0%	94.3%
Aug. 20	**96.0%**	**96.6%**

Table 28: CHRP Examination Validation Committee – Pass mark approval

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.

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 HRPA.

Table 29 provides the means and standard deviations for the domains and for the total score, using all candidates who took the August 2020 CHRP ELE. Table 30 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. Figure 3 shows the distribution of scores for all candidates, along with the pass mark.

Table 29: Total and domain scores for all candidates

	Domain	Percentage	Mean	SD*
A	Employment Contracts and Terminations	73%	33.0	4.8
B	Employer Obligations	69%	23.5	3.6
C	Regulations and Legislation	67%	14.1	2.6
	Total score	70.6%	70.6	9.4

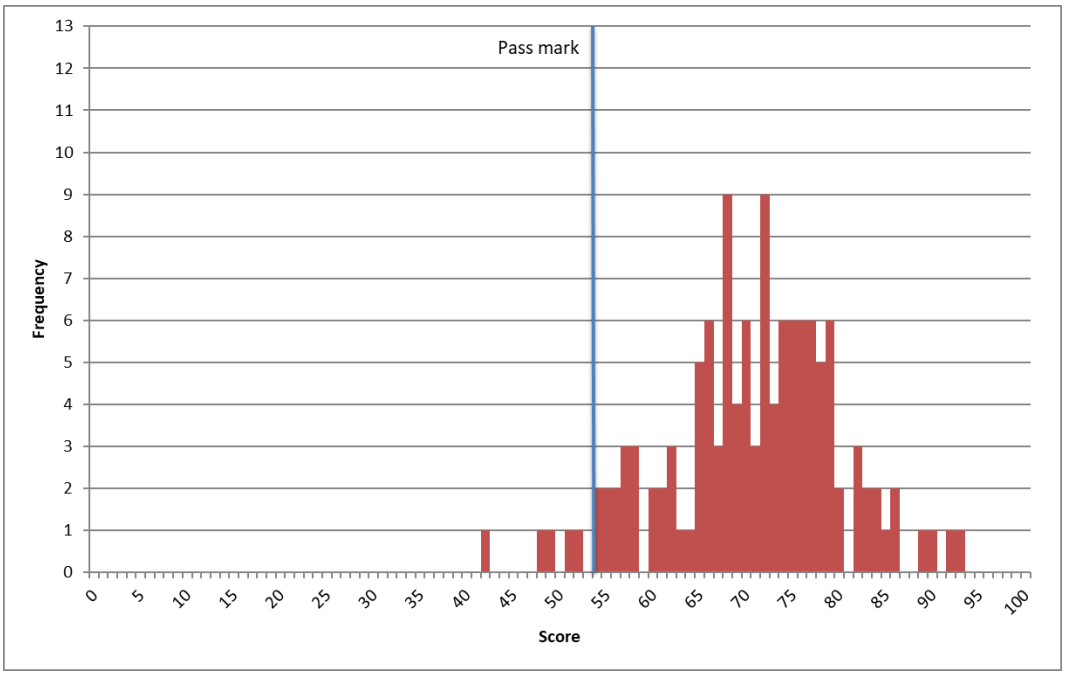
*SD = Standard deviation.

Table 30: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.58	.65
B			.48
C			

*See Table 29 for the full name of each functional area.

Figure 3: Score distribution for all candidates



Key Examination Metrics

Table 31 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 31: Key examination metrics – Candidates included in analysis only

Index	August 2020	January 2020	September 2019	May 2019	January 2019
Scored items	100	102	100	102	101
Candidates	116	122	98	158	119
Mean	71.53 (71.5%)	72.42 (71.0%)	72.69 (72.7%)	74.62 (73.2%)	72.29 (71.6%)
Median	72 (72.0%)	73 (71.6%)	74 (74.0%)	75 (73.5%)	73 (72.3%)
Skewness	−0.306	−0.212	−0.259	−0.369	0.005
Kurtosis	0.284	−0.532	−0.528	−0.122	−0.244
Range	48–93 (48.0–93.0%)	53–92 (52.0–90.2%)	52–90 (52.0–90.0%)	51–95 (50.0–93.1%)	51–91 (50.5–90.1%)
Standard deviation	8.90	8.80	8.66	9.05	8.01
Cronbach's alpha	.78	.77	.77	.80	.74
Mean r_{pb}*	.17	.16	.16	.18	.15
SEM[i]	4.13	4.19	4.13	4.08	4.11
SEM at the pass mark	4.66	4.67	4.68	4.68	4.59
Decision consistency (uncorrected)[ii]	.95	.93	.94	.96	.95
Perceived fairness[iii]	56%	54%	55%	57%	63%
Pass mark	53.108	56.132	56.721	55.971	56.857
Effective pass mark	54	57	57	56	57
Pass rate	96.6%	94.3%	95.9%	96.8%	98.3%

[i]SEM = standard error of measurement.

[ii]Subkoviac method.

[iii]Based on responses to the post-examination survey for all candidates.

Related Development Activities

Since the last administration of the CHRP ELE in January 2020, the following exam development activities have been completed.

Validation

To renew the validation of items expiring from usability, a validation session was held with the EVC (see Table 32) remotely on April 22, 23, 27 and 29, 2020. Though originally planned to be an in person session, it was converted to remote because of COVID-19. During these 4 sessions, CKE 1 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 32: 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
✓ **Alyssa Young**	CHRL	5–9	2017	Non-profit

✓Participated in the session.

The EVC members received advance materials outlining:

- Purpose of the session
- Description of the CHRP credential
- CHRP ELE blueprint

- Criteria for good test items
- Validation process
- Relevant legislation

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, they 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 two years to make decisions about who would be certified as CHRP
- Move the item to the CHRL 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 209 items and case texts as suitable for the CHRP ELE, rejected 7 items, and shifted 0 items for eligibility in the CHRL ELE bank. Thirty-nine items/case texts were revised prior to validation as part of this exercise. The committee also verified the topic and cognitive level for all items, and added rationales and references where missing, incomplete, or not current.

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 33.

Table 33: 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 34.

Table 34: 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 *Employment Standards Act, 2000*	5%
		C1.2 Probation period under *Employment Standards Act, 2000*	2%
		C1.3 Different leaves permitted under the *Employment Standards Act, 2000*	2%
		C1.4 Employers' obligations under *Employment Standards Act, 2000*	1%

4%	**C2. Occupational Health and Safety Act**	
	C2.1 Making policies that are compliant with the *Occupational Health and Safety Act, 1990*	2%
	C2.2 Ministry of Labour's rights under the *Occupational Health and Safety Act, 1990*	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 *Pay Equity Act, 1990*	2%
1%	**C5. Canada Labour Code**	
	C5.1 Employers' obligations under *Canada Labour Code, 1985*	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 35.

Table 35: 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.

1. Where scenarios or test items include a workplace, the workplace allocation will be as follows:
    a. For profit enterprise, 60% (+/- 10%)
    b. Government, 20% (+/- 5%)
    c. Not-for-profit, 20% (+/- 5%)
2. 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.