

## **Technical Report: May 2019 CHRP ELE**

---

### **Human Resources Professionals Association**

10 June 2019

Version 1.1



# Contents

<b>Executive Summary</b> .....	<b>5</b>
<b>Administration</b> .....	<b>6</b>
Form Setting .....	6
Testing Window .....	7
<b>Analysis</b> .....	<b>8</b>
Data Cleaning and Integrity Checks.....	8
Post-Examination Survey.....	10
Initial Analysis .....	13
Key Validation.....	13
Establishing the Pass Mark: Equating .....	16
Scoring .....	24
<b>Key Examination Metrics</b> .....	<b>26</b>
<b>Related Development Activities</b> .....	<b>27</b>
Item Validation .....	27
<b>Appendix</b> .....	<b>29</b>
Blueprint .....	29

## 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 January 2019 .....	17
Table 15: Equating parameter table – To January 2019.....	17
Table 16: Equating outcome table – To January 2019 .....	18
Table 17: Anchor item fit to blueprint – To May 2018 .....	19
Table 18: Equating parameter table – To May 2018.....	19
Table 19: Equating outcome table – To May 2018 .....	20
Table 20: Anchor item fit to blueprint – To September 2018.....	21
Table 21: Equating parameter table – To September 2018 .....	21
Table 22: Equating outcome table – To September 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.....	25
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: Ad hoc CHRP Examination Validation Committee members – Item validation.....	28

Table 31: CHRP Employment Law Examination Blueprint Structural Variables.....29  
Table 32: CHRP Employment Law Examination Blueprint Content Weights .....31  
Table 33: CHRP Employment Law Examination Blueprint Cognitive Level Weights.....32

## 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 Summary<sup>1</sup>

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 169 candidates using computer-based testing at Prometric test centres May 6–21, 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<sup>2</sup> ( $n=158$ ) was 74.6 (73.2%), and for all candidates it was 73.9 (72.5%), out of 102 validated items for scoring. Reliability was marginal at .80 (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 May 2018, September 2018 and January 2019 administrations, yielding an integer pass mark of 56. 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 95.9% and a pass rate for all candidates of 96.8%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards<sup>3</sup> and ISO 17024 standards.<sup>4</sup>

---

<sup>1</sup> 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.

<sup>2</sup> 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.

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

<sup>4</sup> 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 Kriss Stone (CHRP Examination Validation Committee member) in a remote session held March 11, 2019.

The final form composition for the May 2019 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
<b>A</b> Employment Contracts and Terminations	50	46% ± 5%	46–56	—
<b>B</b> Employer Obligations	37	33% ± 4%	32–40	—
<b>C</b> Regulations and Legislation	23	21% ± 3%	20–26	—
<b>TOTAL</b>	<b>110</b>		<b>110</b>	—

Table 2: Cognitive level fit at administration

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	12	10% ± 3%	8–14	—
Application	62	60% ± 10%	55–77	—
Critical thinking	36	30% ± 10%	22–44	—
<b>TOTAL</b>	<b>110</b>		<b>110</b>	<b>—</b>

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 May 6–21, 2019, inclusive, and 169 candidates wrote the exam.

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 February 22, 2019.

# 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, 58 minutes (up 3 minutes from January 2019). The time limit on the CHRP ELE was 3½ hours, suggesting that time may have been a factor for at least some candidates.

Twenty-two candidates (13%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 9 candidates (5%) 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 February 2019 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 -.03, suggesting that time constraints were not generally related to candidate performance. (Note that one candidate exceeded the time limit; this candidate was 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 increase of 2.3 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

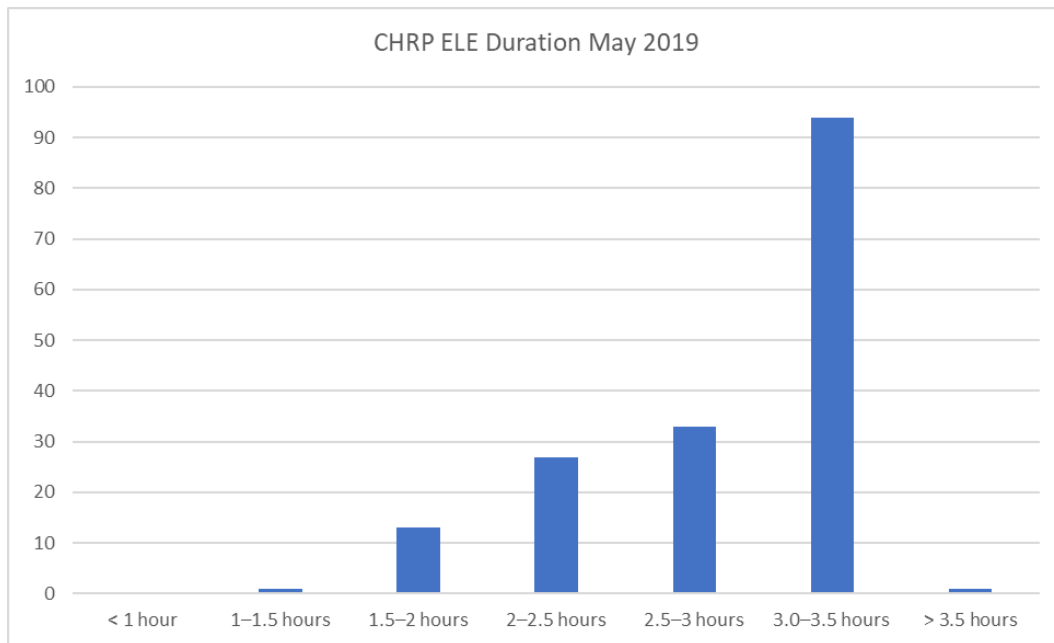
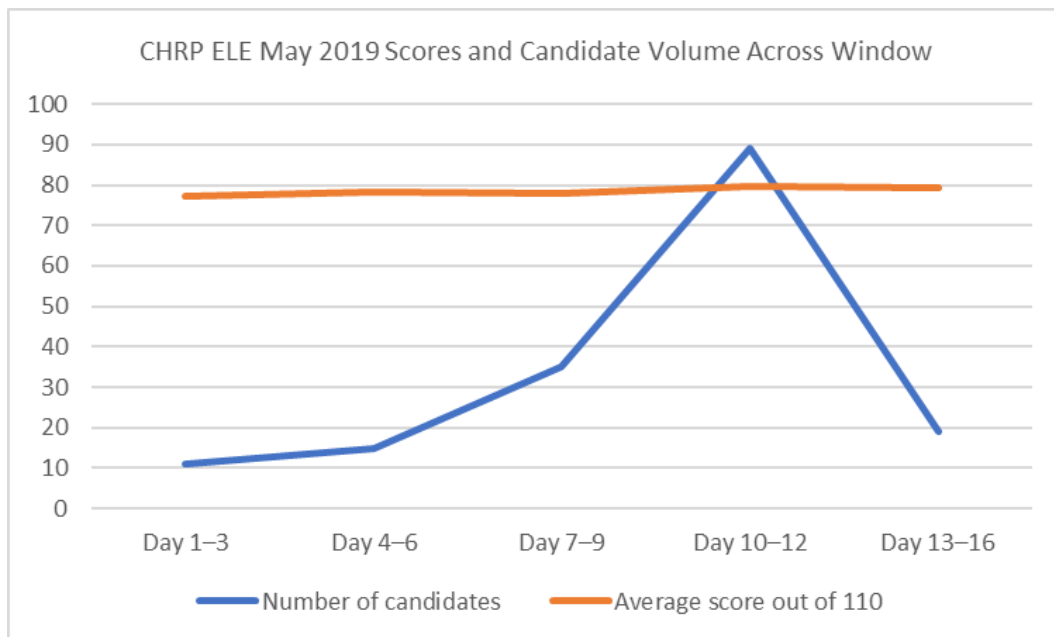


Figure 2: Candidate volume and score trends across testing window



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, 2 candidates were removed from analyses.

Candidates who had failed a previous employment law examination (CHRP ELE or CHRL ELE;  $n=9$ ) scored lower than did those who had not (67.9 and 79.7, respectively, on the full exam of 110 items;  $t(9)=3.74$ ,  $p<.01$ ). In keeping with standard procedures, these candidates were removed from subsequent analyses. The CHRP ELE analysis proceeded with 158 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; 164 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.	61	68	10	20	5	3.98	79%
2.	I was well informed about what documents to bring to the exam location.	97	56	3	5	1	4.50	94%
3.	Proctors enforced the exam-day rules and the security procedures at the test centre were what I expected.	105	45	5	3	3	4.53	93%
4.	Proctors were professional and courteous.	108	38	7	5	1	4.55	92%
5.	The tutorial helped me understand how to complete the examination on the computer.	87	67	5	0	0	4.52	97%
6.	The legislation and case texts were easy to access during the examination.	56	82	13	8	0	4.17	87%
7.	Navigation through the examination was easy and intuitive.	67	75	10	2	1	4.32	92%

\*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.	51	69	15	19	1	3.97	77%
9.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	35	69	25	21	5	3.70	67%
10.	I feel I was adequately prepared to write this examination.	20	81	41	11	2	3.68	65%
11.	The questions in the examination were clearly written.	24	83	33	15	0	3.75	69%
12.	The terminology used in the examination was accurate.	28	107	14	6	0	4.01	87%
13.	The situations presented in the examination were realistic.	42	94	15	3	0	4.14	88%
14.	The questions in the examination reflected the Employment Law Examination blueprint.	21	86	37	9	0	3.78	70%
15.	The examination was a fair assessment of my ability.	17	70	48	16	1	3.57	57%

\*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 case-based 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.	48	32%
I preferred the case-based items.	68	45%
I had no preference between independent and case-based items.	36	24%

**Table 6: Preference regarding number of response options**

	Count	%
I preferred having 3 options.	116	76%
I preferred having 4 options.	7	5%
It did not matter to me how many options were used.	29	19%

**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.	38	25%
I feel that completing the examination on a computer decreased my performance.	20	13%
I feel that completing the examination on a computer had no effect on my performance.	94	62%

**Table 8: Value of access to legislation**

	Count	%
Yes, it was essential to me in completing the examination.	62	41%
Yes, but I only consulted it a few times.	66	43%
No, I could not find the answers to questions I had.	10	7%
No, I did not need to consult it to complete the examination.	9	6%
No, it was more of a distraction than an aid.	5	3%

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	169
Candidates in analysis	158
Mean	79.7 (72.5%)
Standard deviation	9.05
Range	54–101 (49.1–91.8%)
Cronbach's alpha	.78
Disattenuated alpha	.85
Mean $r_{pb}^*$	.16

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 May 27, 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	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	Health services
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
Ileean 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 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. 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 8 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 102 scored items was .80. The difficulties ranged from 31.6% to 97.5%, with a mean of 73.2%. The  $r_{pb}^*$  values ranged from  $-.06$  to  $.43$ , with a mean of  $.18$ . 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 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 approved the final set of items for use in scoring the May 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	47	46% ± 5%	42–52	—
B Employer Obligations	33	33% ± 4%	30–37	—
C Regulations and Legislation	22	21% ± 3%	19–24	—
<b>TOTAL</b>	<b>102</b>		<b>102</b>	<b>—</b>

Table 12: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	12	10% ± 3%	8–13	—
Application	58	60% ± 10%	51–71	—
Critical thinking	32	30% ± 10%	21–40	—
<b>TOTAL</b>	<b>102</b>		<b>102</b>	<b>—</b>

Table 13: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	28	25% ± 3%	23–28	—
Case	74	75% ± 3%	74–79	—
<b>TOTAL</b>	<b>102</b>		<b>102</b>	<b>—</b>

## Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014)<sup>5</sup> and Livingston and Kim (2009),<sup>6</sup> was used to establish the pass mark for the May 2019 CHRP ELE. The goal of this process was to set a pass mark for the May 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 (May 2018, September 2018 and January 2019). The intention following these 3 equating runs was to average them to arrive at a final pass mark for the May 2019 CHRP ELE. These administrations were chosen as the most recent administration (January 2019), the administration from 1 year in the past (May 2018, and a third recent administration (September 2018) because of variance in the first two equating runs.

### Equating Back to the January 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; 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 0.73 and a mean corrected point-biserial of .22.

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.

---

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

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



Table 14: Anchor item fit to blueprint – To January 2019

	Area	Actual	Target
<b>A</b>	Employment Contracts and Terminations	42%	46%
<b>B</b>	Employer Obligations	36%	33%
<b>C</b>	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 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 May 2019 CHRP ELE scored higher than the population taking the January 2019 CHRP ELE (73.5% vs. 72.0%, respectively;  $t(275)=1.02$ , *ns*). Because the May 2019 CHRP ELE candidates were of greater ability (based on the anchors, non-significance notwithstanding), they should have a higher pass rate (though with small samples the tails of the distribution will not be regular).

The equating analysis generally did not bear this out (Table 16), though this was due to more candidates at the lower tail of the distribution. The methods show a pass mark of 57 to 59. Given the sample sizes involved, Tucker would be the primary method under consideration and the equated value of 57.36 was carried forward in the analysis.

Table 15: Equating parameter table – To January 2019

		Jan. 2019	May 2019
	n	119	158
	Scored items	101	102
Mean score	Total	71.6%	73.2%
	Anchors	72.0%	73.5%

Table 16: Equating outcome table – To January 2019

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equated Jan. 2019	56.86	57	97.0%	98.3%
Tucker	57.36	58	93.5%	95.6%
Levine observed	57.95	58	93.5%	95.6%
Mean	58.29	59	92.9%	94.9%
Circle Arc 1	57.83	58	93.5%	95.6%
Circle Arc 2	57.82	58	93.5%	95.6%
Braun-Holland	56.22	57	94.7%	96.2%

### 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; 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 0.74 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 aligned with the targets and reflect the scope and weighting across the full exam.

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

	Area	Actual	Target
<b>A</b>	Employment Contracts and Terminations	47%	46%
<b>B</b>	Employer Obligations	33%	33%
<b>C</b>	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 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 May 2019 CHRP ELE scored higher than the population taking the May 2018 CHRP ELE (73.9% vs. 72.6%, respectively;  $t(269)=0.90$ , *ns*). Because the May 2019 CHRP ELE candidates were of higher ability (based on the anchors, non-significance notwithstanding), they should have a higher pass rate (though, again, the tails of the distribution will be more erratic with small samples).

The equating analysis approximates this, with the pass rate being up marginally (Table 19). All methods show a pass mark of 55 to 58. 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 May 2018

		May 2018	May 2019
	n	113	158
	Scored items	100	102
Mean score	Total	71.3%	73.2%
	Anchors	72.6%	73.9%

Table 19: Equating outcome table – To May 2018

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
<b>Equated May 2018</b>	<b>54.85</b>	<b>55</b>	<b>97.5%</b>	<b>97.3%</b>
Tucker	54.59	55	96.4%	97.5%
Levine observed	54.21	55	96.4%	97.5%
Mean	57.37	58	93.5%	95.6%
Circle Arc 1	56.84	57	94.7%	96.2%
Circle Arc 2	56.83	57	94.7%	96.2%
Braun-Holland	54.22	55	96.4%	97.5%

Because the two planned equating runs showed a difference in pass marks of 2.6 marks, a third equating run was conducted. Note that this divergence is likely explained by differences in variance across administrations.

### 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 0.75 and a mean corrected point-biserial of .21.

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

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

	Area	Actual	Target
<b>A</b>	Employment Contracts and Terminations	47%	46%
<b>B</b>	Employer Obligations	30%	33%
<b>C</b>	Regulations and Legislation	23%	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 (74.7% vs. 72.4%, respectively;  $t(261)=1.42$ , *ns*). Because the May 2019 CHRP ELE candidates were of higher ability (based on the anchors, non-significance notwithstanding), they should have a higher 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 shows a higher pass rate (Table 22) but also a lot of divergence in equated values. The methods show a pass mark range of 52 to 58. 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 September 2018

		Sep. 2018	May 2019
	N	105	158
	Scored items	98	102
Mean score	Total	71.7%	73.2%
	Anchors	72.4%	74.7%

Table 22: 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	53.93	54	96.4%	97.5%
Levine observed	51.77	52	98.8%	98.7%
Mean	57.78	58	93.5%	95.6%
Circle Arc 1	56.65	57	94.7%	96.2%
Circle Arc 2	56.65	57	94.7%	96.2%
Braun-Holland	54.01	55	96.4%	97.5%

### 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 the September equating run are clearly more variable than the other two runs. 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 (55.971) of the January 2019 and May 2018 equating runs.

Note that the decision to remove the September administration from consideration was also taken when analyzing the January 2019 examination. Furthermore, if the September result was included in the average, the same integer pass mark (56) would have resulted.

Using the established convention for this testing program, the combined pass mark would be rounded up to a cut score of 56. The resulting pass rate for first-time candidates (96.8%) is slightly lower than that observed for the May 2018 and January 2019 CHRP ELE: 97.3% and 98.3%, respectively. The pass rate for all candidates was 95.9% which was also just slightly lower than the 97.5% in May 2018 and the 97.0% in January 2019. 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 May 30, 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 (via the Associate Registrar) accepted the recommended pass mark and so the pass mark was formally established.

Table 23: Equating outcome table – Combined results

	May 18	Sep. 18	Jan. 19
Tucker	54.6	53.9	57.4
Levine observed	54.2	51.8	57.9
Mean	57.4	57.8	58.3
Circle arc 1	56.8	56.7	57.8
Circle arc 2	56.8	56.7	57.8
Braun-Holland	54.2	54.0	56.2

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%
May 19	<b>95.9%</b>	<b>96.8%</b>

Table 25: 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	Health services
✓ 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.

## 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 26 provides the means and standard deviations for the domains and for the total score, using all candidates who took the May 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%	33.4	5.0
B Employer Obligations	74%	24.4	3.5
C Regulations and Legislation	73%	16.1	2.8
<b>Total score</b>	<b>72.5%</b>	<b>73.9</b>	<b>9.4</b>

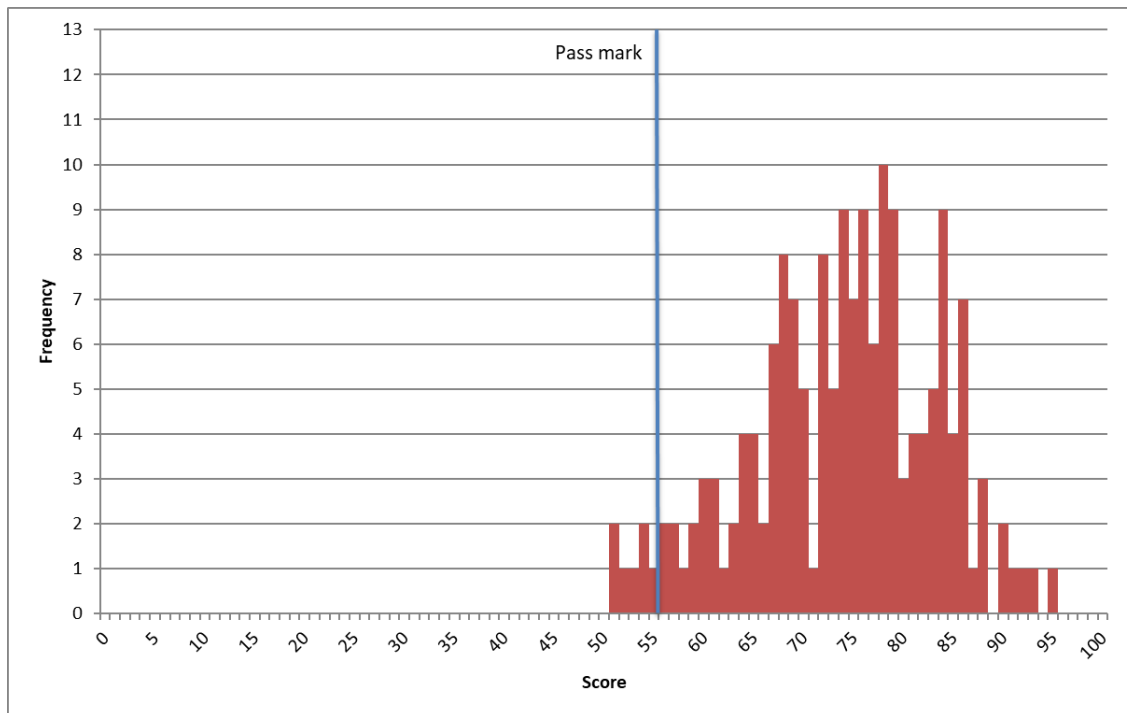
\*SD = Standard deviation.

Table 27: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.55	.49
B			.46
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	May 2019	January 2019	September 2018	May 2018	January 2018
Scored items	102	101	98	100	100
Candidates	158	119	105	113	111
Mean	74.62 (73.2%)	72.29 (71.6%)	70.30 (71.7%)	71.30 (71.3%)	73.03 (73.0%)
Median	75 (73.5%)	73 (72.3%)	72 (73.5%)	71 (71.0%)	73 (73.0%)
Skewness	-0.369	0.005	-0.593	-0.350	-0.432
Kurtosis	-0.122	-0.244	-0.121	0.281	-0.023
Range	51–95 (50.0– 93.1%)	51–91 (50.5– 90.1%)	49–84 (50.0– 85.7%)	50–88 (50.0– 88.0%)	49–90 (49.0– 90.0%)
Standard deviation	9.05	8.01	8.04	7.80	8.22
Cronbach's alpha	.80	.74	.75	.72	.76
Mean $r_{pb}^*$	.18	.15	.15	.14	.15
SEM <sup>i</sup>	4.08	4.11	4.06	4.09	4.04
SEM at the pass mark	4.68	4.59	4.55	4.57	4.59
Decision consistency (uncorrected) <sup>ii</sup>	.96	.95	.94	.95	.95
Perceived fairness <sup>iii</sup>	57%	63%	64%	47%	59%
Pass mark	55.971	56.857	54.718	54.846	56.817
Effective pass mark	56	57	55	55	57
Pass rate	96.8%	98.3%	95.2%	97.3%	97.3%

<sup>i</sup>SEM = standard error of measurement.

<sup>ii</sup>Subkoviak method.

<sup>iii</sup>Based on responses to the post-examination survey for all candidates.

## Related Development Activities

Since the last administration of the CHRP ELE in January 2019, the following exam development activities have taken place.

### Item Validation

To provide sufficient items for upcoming administrations, a 2½-day validation session was held February 20–22, 2019, at HRPAs offices. The CHRP EVC members who participated are listed in Table 29 (one of the available members was not available on the last day). To augment the EVC members available for this session, ad hoc members were added, as listed in Table 30 (one of the adhoc members was unable to continue after day 1). 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	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	Health services
Tanya Gopaul	CHRL	10–15	2017	Banking
✓ Jean Lazarus	CHRL	15–19	2017	Health services
Suman Seth	CHRL	15–19	2018	Government
Kriss Stone	CHRL	10–15	2017	Real estate
Ilelean Tait	CHRL	15–20	2017	Environmental
✓ Karen Weiler	CHRL	20–29	2017	Software/ communications
Alyssa Young	CHRL	5–9	2017	Non-profit

✓ Participated in the session.

Table 30: Ad hoc CHRP Examination Validation Committee members – Item validation

Member	Credential	Years of Relevant Experience	Industry
✓ Florence Aromolaran	CHRL, BA	10–15	Non-profit
✓ Portia Daisy	CHRL, BFA, PMP	10–15	Education
✓ Kelly Gillis	CHRL	10–15	Healthcare
✓ Karen Pantaleo	CHRL, BBA	20–29	Consultant
✓ Patricia Verkley	CHRL, BA	10–15	Non-profit

✓ Participated in the 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, full training was provided to inform the ad hoc committee members of the expectations for the CHRP credential and the process used to validate items.

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 CKE 1 or 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 12–20 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.

The panel members reviewed and validated 145 items as suitable for the CHRP ELE, designated 6 items to move to the CHRL ELE bank, designated 1 items to move to the CKE 1 bank, and rejected zero (0) items as unfixable. Thirty-nine of the 145 items were edited prior to being validated.

# 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 HRP A 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 31.

**Table 31: 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 32.

Table 32: CHRP Employment Law Examination Blueprint Content Weights

Category Weight	Topic Weight	Topic	Subtopic Weight
<b>46%</b>	<b>A.</b>	<b>Employment Contracts and Terminations</b>	
	<b>28%</b>	<b>A1. Termination</b>	
		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%
	<b>11%</b>	<b>A2. Contracts</b>	
		A2.1 Contracts and employment agreements	9%
		A2.2 Collective bargaining contracts	2%
	<b>7%</b>	<b>A3. Employee Benefits and Perquisites</b>	
		A3.1 Vacation time, vacation pay and bonuses	5%
		A3.2 Overtime exemptions	2%
<b>33%</b>	<b>B.</b>	<b>Employer Obligations</b>	
	<b>16%</b>	<b>B1. Duty to Accommodate</b>	
		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%
	<b>9%</b>	<b>B2. Misconduct in the Workplace</b>	
		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%
	<b>6%</b>	<b>B3. Common Law</b>	
		B3.1 Including consideration of Common Law principles	5%
		B3.2 Employers' obligations under Common Law	1%
	<b>2%</b>	<b>B4. Sale of Business</b>	
		B4.1 The effects of the sale of the business	2%
<b>21%</b>	<b>C.</b>	<b>Regulations and Legislation</b>	
	<b>10%</b>	<b>C1. Employment Standards Act</b>	
		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%

<b>4%</b>	<b>C2. Occupational Health and Safety Act</b>	
	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%
<b>4%</b>	<b>C3. Jurisdiction</b>	
	C3.1 The difference between federal and provincial legislations	2%
	C3.2 Determining governing legislation when the organization is interprovincial	2%
<b>2%</b>	<b>C4. Pay Equity Act</b>	
	C4.1 Application of <i>Pay Equity Act, 1990</i>	2%
<b>1%</b>	<b>C5. Canada Labour Code</b>	
	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 33.

Table 33: 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.