

Technical Report: September 2019 CHRL ELE

Human Resources Professionals Association

11 October 2019



Contents

Executive Summary	5
Administration	6
Form Setting	6
Testing Window	7
Analysis	8
Data Cleaning and Integrity Checks.....	8
Post-Examination Survey.....	10
Initial Analysis	13
Key Validation.....	14
Establishing the Pass Mark: Equating.....	16
Scoring	24
Key Examination Metrics	26
Related Development Activities	27
Appendix	28
Blueprint	28

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: CHRL 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.....	17
Table 15: Equating parameter table	17
Table 16: Equating outcome table.....	18
Table 17: Anchor item fit to blueprint.....	19
Table 18: Equating parameter table	19
Table 19: Equating outcome table.....	20
Table 20: Anchor item fit to blueprint.....	21
Table 21: Equating parameter table	21
Table 22: Equating outcome table.....	22
Table 23: Equating outcome table – Combined results	23
Table 24: Historical pass rates	23
Table 25: CHRL 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: Employment Law Examination Blueprint Structural Variables.....	28
Table 30: Employment Law Examination Blueprint Content Weights	30
Table 31: CHRL Employment Law Examination Blueprint Cognitive Level Weights	31

List of Figures

Figure 1: Examination time distribution for all candidates.....	9
Figure 2: Candidate volume and score trends across testing window	9
Figure 3: Score distribution for all candidates.....	25

Executive Summary¹

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

The CHRL Employment Law Exam (CHRL ELE) was administered to 200 candidates using computer-based testing at Prometric test centres September 16–30, 2019, inclusive. The examination comprised 110 three-option multiple choice items and had a 3½-hour time limit.

As per the CHRL ELE blueprint, the exam was scored using the 98–102 best-performing items (while adhering to the prescribed distribution across topics). The mean score for first-time candidates ($n=169^2$) was 75.0 (73.5%), and for all candidates it was 74.1 (72.7%), out of 102 validated items for scoring. Reliability was borderline at .75 (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 September 2018, January 2019 and May 2019 administrations, yielding an integer pass mark of 65. 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 CHRL ELE. This pass mark resulted in a pass rate for first-time candidates of 89.3% and a pass rate for all candidates of 87.5%.

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

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

² Excludes those who had failed an HRPAs 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 Laurie Torno (CHRL Examination Validation Committee member), facilitated remotely, in a session held June 19, 2019. A second CHRL EVC member was planned to participate, but due to a travel interruption was not able to.

The final form composition for the primary September 2019 CHRL ELE is shown in Table 1 (domain weighting) and Table 2 (cognitive level weighting). The form reflected the examination blueprint (see Appendix for full CHRL ELE blueprint).

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

Table 1: Domain fit at administration

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	49	46% ± 5%	46–56	—
B Employer Obligations	35	33% ± 4%	32–40	—
C Regulations and Legislation	26	21% ± 3%	20–26	—
TOTAL	110		110	—

Table 2: Cognitive level fit at administration

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	12	10% ± 3%	8–14	—
Application	52	50% ± 10%	44–66	—
Critical thinking	46	40% ± 10%	33–55	—
TOTAL	110		110	—

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

Testing Window

The examination was administered via computer-based testing at Prometric test sites primarily in Ontario. The testing window was September 16–30, 2019, inclusive, and 200 candidates wrote the exam. A technical issue required one candidate to write outside this window and one wrote on paper as an accommodation.

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

Provincial

- AODA – *Accessibility for Ontarians with Disabilities Act, 2005*
- ESA – *Employment Standards Act, 2000*
- LRA – *Labour Relations Act, 1995*
- OHRC – *Human Rights Code*
- OHS Act – *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 June 24, 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 (3 minutes more than in May 2019). The time limit on the CHRL ELE was 3½ hours, suggesting that time may have been a factor for at least some candidates.

Twenty candidates (10%) took the full 3½ hours, suggesting that those candidates may have wanted more time. Five 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 and are close to what was observed at the May 2019 administration; at present they do not appear problematically high. 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 modest at a value of .22, suggesting that time was only modestly related to candidate performance with candidates who took longer getting slightly higher scores.

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 decrease of 2.8 marks out of 110.

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

Figure 1: Examination time distribution for all candidates

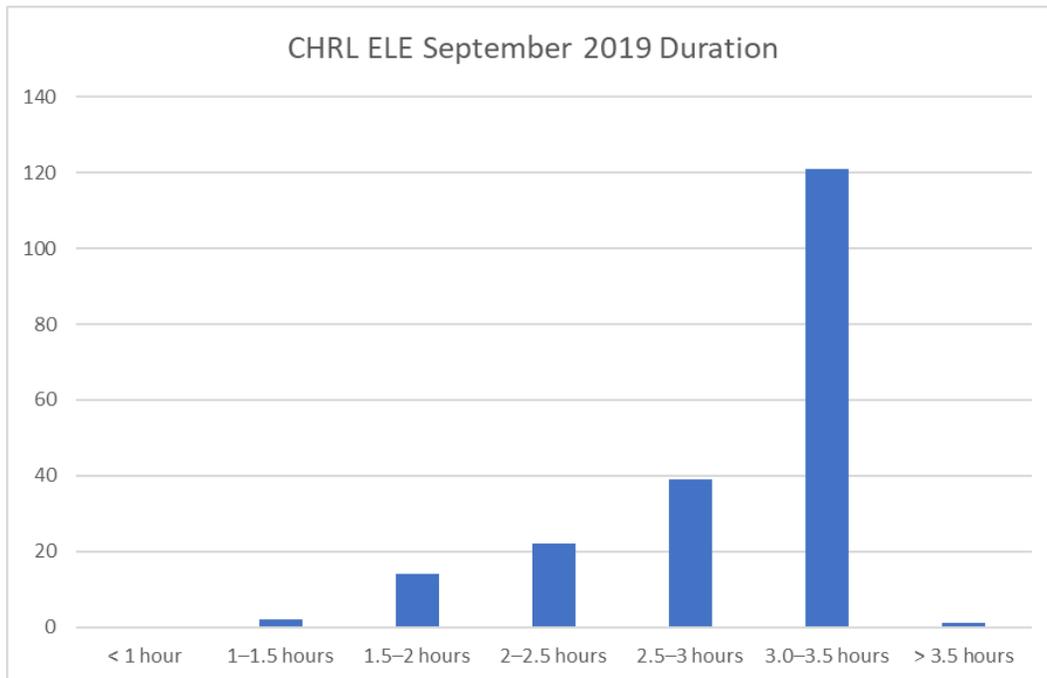
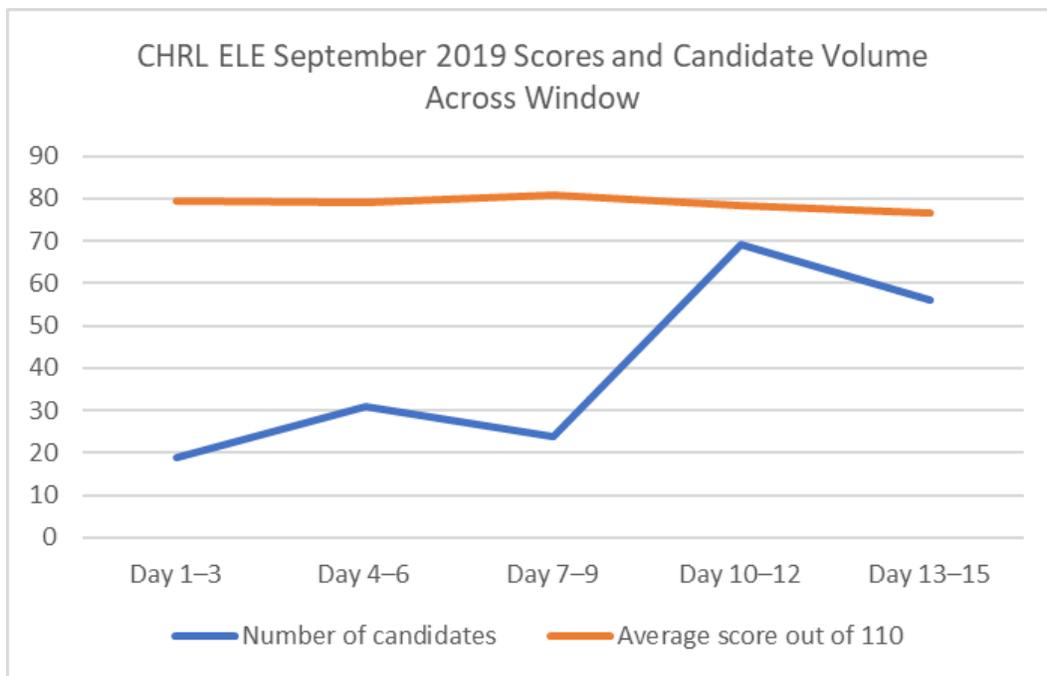


Figure 2: Candidate volume and score trends across testing window



After removing candidates who were administered a previously used test form (who were scored using the same decisions employed at the time that form was originally used), scores were calculated for all remaining candidates based on the full set of 110 items. One candidate was flagged for an abnormally low or high score (z value outside +/- 3.0); however, this

candidate's score was in line with low scores from past administrations and was only marginally outside the threshold and so was left in the analysis. 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. All outliers were removed from initial analyses; candidates with abnormal response patterns (such as having 5 or more blanks) were also 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, 5 candidates were removed from analyses.

Candidates who had failed a previous employment law examination (CHRP ELE or CHRL ELE) scored lower than did those who had not (72.7 and 79.3, respectively, on the full exam of 110 items). This difference was statistically significant ($t(30)=3.40$, $p<.01$), as is typical of repeat test takers. In keeping with standard procedures, all repeat candidates were removed from subsequent analyses. The CHRL ELE analysis proceeded with 169 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 CHRL ELE; 194 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.	102	68	7	10	7	4.28	88%
2.	I was well informed about what documents to bring to the exam location.	121	65	4	2	0	4.59	97%
3.	Proctors enforced the exam-day rules and the security procedures at the test centre were what I expected.	133	52	3	1	1	4.66	97%
4.	Proctors were professional and courteous.	126	51	4	7	2	4.54	93%
5.	The tutorial helped me understand how to complete the examination on the computer.	111	66	10	2	0	4.51	94%
6.	The legislation and case texts were easy to access during the examination.	67	88	15	17	1	4.08	82%
7.	Navigation through the examination was easy and intuitive.	81	88	9	7	2	4.28	90%

*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.	72	73	17	20	5	4.00	78%
9.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	50	81	35	16	5	3.83	70%
10.	I feel I was adequately prepared to write this examination.	21	98	50	16	2	3.64	64%
11.	The questions in the examination were clearly written.	22	105	35	22	3	3.65	68%
12.	The terminology used in the examination was accurate.	29	124	28	6	0	3.94	82%
13.	The situations presented in the examination were realistic.	32	127	18	9	1	3.96	85%
14.	The questions in the examination reflected the Employment Law Examination blueprint.	31	96	42	15	3	3.73	68%
15.	The examination was a fair assessment of my ability.	20	90	49	21	7	3.51	59%

*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 no overall preference between case-based and independent items but a strong preference for 3-option multiple choice items. Most indicated that taking the test on a computer likely improved or had no effect on their performance. 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.	57	31%
I preferred the case-based items.	66	35%
I had no preference between independent and case-based items.	63	34%

Table 6: Preference regarding number of response options

	Count	%
I preferred having 3 options.	149	80%
I preferred having 4 options.	9	5%
It did not matter to me how many options were used.	28	15%

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

	Count	%
I feel that completing the examination on a computer improved my performance.	63	34%
I feel that completing the examination on a computer decreased my performance.	28	15%
I feel that completing the examination on a computer had no effect on my performance.	95	51%

Table 8: Value of access to legislation

	Count	%
Yes, it was essential to me in completing the examination.	71	38%
Yes, but I only consulted it a few times.	79	42%
No, I could not find the answers to questions I had.	27	15%
No, I did not need to consult it to complete the examination.	5	3%
No, it was more of a distraction than an aid.	4	2%

An open-ended question was also posed to candidates asking for any additional comments. Those comments were provided to HRP A for information and consideration. Nothing actionable with respect to scoring emerged in these comments.

Initial Analysis

The full CHRL ELE examination was 110 items, of which approximately 100 were to be scored. The other 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	CHRL ELE
Items	110
Total candidates	200
Candidates in analysis	169
Mean	79.4 (72.2%)
Range	55–97 (50.0–88.2%)
Standard deviation	8.10
Cronbach's alpha	.74
Disattenuated alpha	.86
Mean r_{pb}^*	.14

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 at the extremes of difficulty. Because of the relatively low variance and sample size, items with marginally negative point-biserials were to be expected and these low values were not necessarily indicative of poor item quality. Items were ranked from worst performing to best performing accordingly.

Key Validation

Key validation was conducted via web meeting on October 8, 2019, using members of the CHRL 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 CHRL ELE.

Table 10: CHRL Examination Validation Committee – Key validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓ Jennifer Borges	CHRL	10–14	2017	Manufacturing
Annette Dhanasar	CHRL	15–19	2017	Technology
Debbie Brandt	CHRL	10–14	2017	Government and public centre agencies
Christine Kelsey	CHRL	5–9	2017	Entertainment
✓ Jennifer King	CHRL	20–29	2017	Business and professional services
Nancy Richard	CHRL	15–19	2017	Regulation/Public sector
✓ Kristin Rivait	CHRL	15–19	2017	Health care
✓ Lisa Scian	CHRL	15–19	2017	Information & communication technology
Laurie Torno	CHRL	20–29	2018	Post-secondary education

✓ Participated in the session.

The group was informed that test reliability, as measured by Cronbach's alpha, was .74 based on the set of 110 potentially scored items and that this was below the generally accepted threshold of .80. The group was reminded that restriction of range attenuates this 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. One item was also flagged for review by the group based solely on candidate comments. 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 based purely on the statistics, 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

.75. The difficulties ranged from 30.8% to 97.0%, with a mean of 73.5%. The r_{pb}^* values ranged from $-.18$ to $.41$, with a mean of $.15$.

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

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

Table 11: Domain fit for final scored items

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	46	46% ± 5%	42–52	—
B Employer Obligations	33	33% ± 4%	30–37	—
C Regulations and Legislation	23	21% ± 3%	19–24	—
TOTAL	102		102	—

Table 12: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	11	10% ± 3%	8–13	—
Application	47	50% ± 10%	41–61	—
Critical thinking	44	40% ± 10%	31–51	—
TOTAL	102		102	—

Table 13: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	23	25% ± 3%	23–28	—
Case	79	75% ± 3%	74–79	—
TOTAL	98		98	—

Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014),⁵ was used to establish the pass mark for the September 2019 CHRL ELE. The goal of this process was to set a pass mark for the September 2019 CHRL ELE that would be equivalent to that set for past 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 CHRL ELE was last set after the January 2018 offering of the CHRL 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 (September 2018, January 2019, and May 2019). The intention following these equating runs was to average them to arrive at a final pass mark for the September 2019 CHRL ELE.

Equating Back to the May 2019 Administration

Multiple equating methods were considered for setting the pass mark and these analyses were conducted once key validation was complete. Linear equating was the primary method under considering owing to both samples having 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 .18.

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

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

Table 14: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	48%	46%
B	Employer Obligations	30%	33%
C	Regulations and Legislation	21%	21%

The mean, Tucker, Levine observed-score, circle arc, equipercentile and Braun-Holland methods were computed to ascertain concordance of solutions. Given the sample sizes, similarities of test parameters and lower equating error, the Tucker method 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 sample taking the September 2019 CHRL ELE scored modestly less than the sample taking the May 2019 CHRL ELE (73.5% vs. 74.6%, respectively; $t(345)=1.19$, *ns*). Because the September 2019 CHRL ELE candidates were of modestly lower ability (nonsignificance notwithstanding), they should have modestly lower pass rate.

The equating analysis bears this out, for the most part (Table 16). All methods show a pass mark of 66 or 67, with the recommended solution showing a 66 which leads to pass rate that is marginally higher than seen in May 2019. This minor deviation from expectation is due to sampling fluctuating where the September 2019 score distribution was ‘thinner’ in the left tail (see Figure 3), and so fewer fell under the pass mark.

Table 15: Equating parameter table

		May 2019	Sep. 2019
	N	178	169
	Scored items	98	102
Mean score	Total	74.6%	73.5%
	Anchors	75.5%	74.0%

Table 16: Equating outcome table

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equating May 2019	63.38	64	83.5%	86.0%
Tucker	65.92	66	85.5%	88.2%
Levine observed	65.95	66	85.5%	88.2%
Mean	66.18	67	83.0%	86.4%
Circle Arc 1	65.90	66	85.5%	88.2%
Circle Arc 2	65.90	66	85.5%	88.2%
Equipercentile	65.88	66	85.5%	88.2%
Braun-Holland	65.93	66	85.5%	88.2%

Equating Back to the September 2018 Administration

Multiple equating methods were considered for setting the pass mark and these analyses were conducted once key validation was complete. Linear equating was the primary method under considering owing to both samples having 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 0.73 and a mean corrected point-biserial of .19.

Table 17 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 approximate weighting across the full exam.

Table 17: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	48%	46%
B	Employer Obligations	35%	33%
C	Regulations and Legislation	17%	21%

The mean, Tucker, Levine observed-score, circle arc, equipercentile and Braun-Holland methods were computed to ascertain concordance of solutions. Given the sample sizes, similarities of test parameters and lower equating error, 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 sample taking the September 2019 CHRL ELE scored marginally lower than the sample taking the September 2018 CHRL ELE (72.6% vs. 73.4%, respectively; $t(334)=0.55$, *ns*). Because the September 2019 CHRL ELE candidates were of marginally lower ability (nonsignificance notwithstanding), they should have a marginally lower or the same pass rate.

The equating analysis bears this out in general (Table 19). All methods indicate a pass mark of 64–65, with the recommended Tucker method showing a 65 with a modest increase in the pass rate from 87.4% to 89.3%. As already described, this increase (rather than the expected decrease) is likely due to sampling fluctuation at the low end of scoring.

Table 18: Equating parameter table

		Sep. 2018	Sep. 2019
	N	167	169
	Scored items	99	102
Mean score	Total	74.2%	73.5%
	Anchors	72.6%	73.4%

Table 19: Equating outcome table

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equating Sep. 2018	63.86	64	86.2%	87.4%
Tucker	64.21	65	87.5%	89.3%
Levine observed	63.72	64	89.5%	91.7%
Mean	65.00	65	87.5%	89.3%
Circle Arc 1	64.55	65	87.5%	89.3%
Circle Arc 2	64.54	65	87.5%	89.3%
Equipercentile	64.04	65	87.5%	89.3%
Braun-Holland	63.93	64	89.5%	91.7%

The average of the two Tucker values falls close the integer (65.07), and so a third equating run was conducted to determine the final value.

Equating Back to the January 2019 Administration

Multiple equating methods were considered for setting the pass mark and these analyses were conducted once key validation was complete. Linear equating was the primary method under considering owing to both samples having 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 .17.

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

Table 20: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	47%	46%
B	Employer Obligations	30%	33%
C	Regulations and Legislation	23%	21%

The mean, Tucker, Levine observed-score, circle arc, equipercentile and Braun-Holland methods were computed to ascertain concordance of solutions. Given the sample sizes, similarities of test parameters and lower equating error, the Tucker method 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 sample taking the September 2019 CHRL ELE scored almost the same as the sample taking the January 2019 CHRL ELE (74.9% vs. 74.8%, respectively; $t(344)=0.10$, *ns*). Because the September 2019 CHRL ELE candidates were of the same ability, they should have about the same pass rate.

The equating analysis bears this out, for the most part (Table 16). All methods show a pass mark of 64 or 65, with the recommended solution showing a 65 which leads to pass rate that is modestly higher than seen in January 2019. This minor deviation from expectation is again due to sampling fluctuating.

Table 21: Equating parameter table

		Jan. 2019	Sep. 2019
	N	177	169
	Scored items	98	102
Mean score	Total	73.1%	73.5%
	Anchors	74.9%	74.8%

Table 22: Equating outcome table

Method	Pass Mark		Pass Rate	
	Precise	Integer	All	First Time
Equating Jan. 2019	60.74	61	85.7%	85.9%
Tucker	64.63	65	87.5%	89.3%
Levine observed	63.82	64	89.5%	91.7%
Mean	64.15	65	87.5%	89.3%
Circle Arc 1	63.67	64	89.5%	91.7%
Circle Arc 2	63.67	64	89.5%	91.7%
Equipercntile	64.49	65	87.5%	89.3%
Braun-Holland	64.60	65	87.5%	89.3%

Combined Results

Table 23 shows the pass mark values across the 3 equating runs. The value highlighted in green is the one that would be selected based on sample parameters at each equating run. Overall, the difference between the Tucker values is small, and the optimal solution without clear reason to do otherwise is to take the weighted average across the 3 values (equalling 64.974 out of 102). Weighting was done by number of anchor items and number of candidates in the previous administration (on the premise that more anchors and more candidates produce a more stable solution). This procedure should serve to lessen the impact of sample-dependent fluctuations.

Using the established convention for this testing program, the mean value would be rounded up to a cut score of 65. The resulting pass rate for first-time candidates (89.3%) is comparable to the values observed on previous administrations, as would be expected barring more extreme changes in candidate performance. It is modestly higher, but again this is simply a matter of score fluctuation at the lower end of performance. The pass rate for all candidates (87.5%) was also similar to that seen in recent administrations. See Table 24 for historical pass rates.

The final pass mark value, and the process used to derive it, was presented to the CHRL EVC (Table 25) via teleconference on October 11, 2019. No concerns were raised regarding the pass mark or pass rate, nor regarding the method used to finalize the value. The panel formally approved the pass mark (which was presented along with the consequent pass rate data) for recommendation to HRP. The HRP Associate Registrar approved the panel's recommendation on the call.

Table 23: Equating outcome table – Combined results

	Sep. 18	May 19	Jan. 19
Tucker	64.2	65.9	64.6
Levine observed	63.7	65.9	63.8
Mean	65.0	66.2	64.1
Circle arc 1	64.5	65.9	63.7
Circle arc 2	64.5	65.9	63.7
Equipercentile	64.0	65.9	64.5
Braun-Holland	63.9	65.9	64.6

Table 24: Historical pass rates

	Pass rate	
	All	First-time
Jan. 2017	81.1%	85.2%
May	90.7%	91.2%
Sep.	88.8%	91.0%
Jan. 2018	86.1%	88.6%
May	78.0%	82.7%
Sep.	85.7%	87.4%
Jan. 2019	85.7%	85.9%
May	83.5%	86.0%
Sep.	87.5%	89.3%

Table 25: CHRL Examination Validation Committee – Pass mark approval

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓ Jennifer Borges	CHRL	10–14	2017	Manufacturing
✓ Annette Dhanasar	CHRL	15–19	2017	Technology
Debbie Brandt	CHRL	10–14	2017	Government and public centre agencies
Christine Kelsey	CHRL	5–9	2017	Entertainment
✓ Jennifer King	CHRL	20–29	2017	Business and professional services
Nancy Richard	CHRL	15–19	2017	Regulation/Public sector
Kristin Rivait	CHRL	15–19	2017	Health care
Lisa Scian	CHRL	15–19	2017	Information & communication technology
Laurie Torno	CHRL	20–29	2018	Post-secondary education

✓ Participated in the session.

Scoring

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

Table 26 provides the means and standard deviations for the domains and for the total score, using all candidates who took the September 2019 CHRL 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. 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	74%	34.2	4.7
B Employer Obligations	72%	23.6	3.1
C Regulations and Legislation	71%	16.3	2.7
Total score	72.7%	74.1	8.6

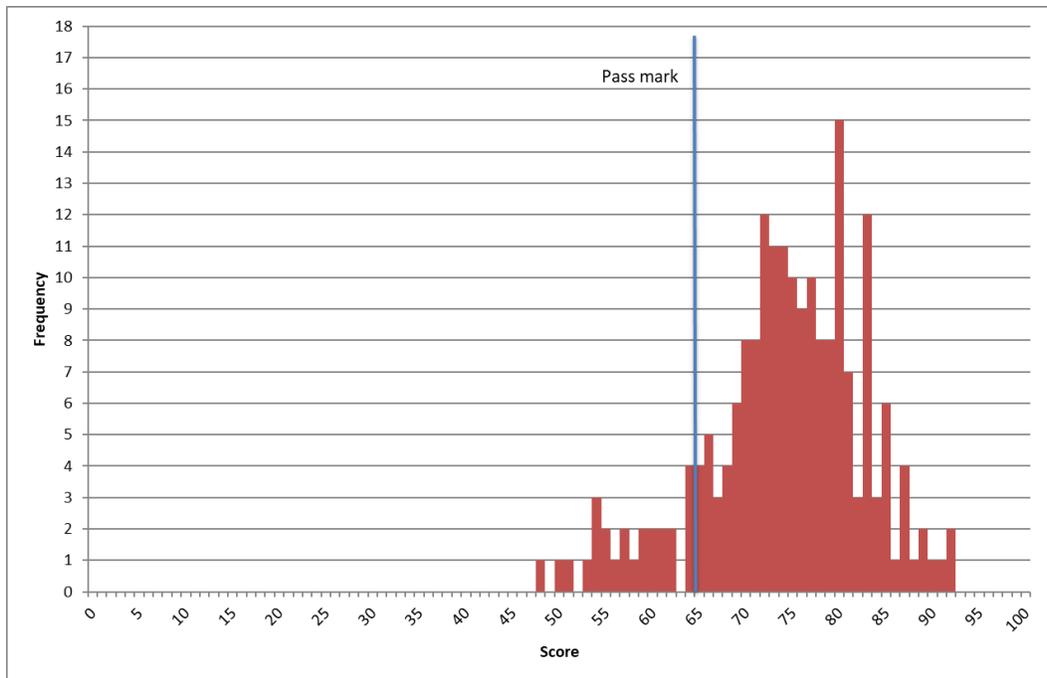
*SD = Standard deviation.

Table 27: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.49	.49
B			.45
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.

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

Index	September 2019	May 2019	January 2019	September 2018	May 2018
Scored items	102	98	98	99	100
Candidates	169	178	177	167 ^{iv}	133
Mean	75.0 (73.5%)	73.1 (74.6%)	71.7 (73.1%)	73.5 (74.2%)	71.5 (71.5%)
Median	76 (74.5%)	74 (75.5%)	73 (74.5%)	75 (75.8%)	71 (71.0%)
Skewness	-0.595	-0.421	-0.469	-0.446	-0.019
Kurtosis	0.486	0.258	-0.212	-0.584	-0.653
Range	50–92 (49.0– 90.2%)	47–91 (48.0– 92.9%)	46–90 (46.9– 91.8%)	53–89 (53.5– 89.9%)	50–91 (50.0– 91.0%)
Standard deviation	8.11	8.20	9.20	7.65	9.36
Cronbach's alpha	.75	.78	.82	.75	.81
Mean r_{pb}^*	.15	.17	.19	.15	.18
SEM ⁱ	4.02	3.89	3.90	3.85	4.05
SEM at the pass mark	4.46	4.35	4.37	4.28	4.45
Decision consistency (uncorrected) ⁱⁱ	.91	.89	.91	.88	.86
Perceived fairness ⁱⁱⁱ	59%	57%	53%	52%	56%
Pass mark	64.974	63.379	60.745	63.859	62.345
Effective pass mark	65	64	61	64	63
Pass rate	89.3%	86.0%	85.9%	87.4%	82.7%

ⁱSEM = standard error of measurement.

ⁱⁱSubkoviac method.

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

^{iv}The original September 2018 analysis was run with one candidate missing from the data set. That candidate has subsequently been added back in.

Related Development Activities

Since the last administration of the CHRL ELE in May 2019, no development activities have taken place.

Appendix

Blueprint

CHRL Employment Law Examination Blueprint

Human Resources Professionals Association

Version 2.0

Approved by CHRL Exam Validation Committee March 13, 2018

Approved by HRP A Registrar March 14, 2018

Effective September 2018 administration

Credential

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

Purpose

The CHRL 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 CHRL level, in Ontario.

Structure

The structural variables provide high level guidance as to what the examination will look like. These appear in Table 29.

Table 29: 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 30.

Table 30: Employment Law Examination Blueprint Content Weights

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

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

Note: Reasonable ranges around the Topic weights are employed.

Cognitive Level

The cognitive level weights are based on Bloom's taxonomy. The purpose of this weighting is generally to ensure that an examination does not unintentionally over-focus on specific types of items, and to provide candidates with a range of items (in approximate proportion) that reflects the cognitive operations they must apply on the job. The weights appear in Table 31.

Table 31: CHRL Employment Law Examination Blueprint Cognitive Level Weights

Level	Weight	Range
Knowledge	10%	+/- 3%
Application	50%	+/- 10%
Critical Thinking	40%	+/- 10%

Miscellaneous Guidance

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

- Where scenarios or test items include a workplace, the workplace allocation will be as follows:
 - For profit enterprise, 60% (+/- 10%)
 - Government, 20% (+/- 5%)
 - Not-for-profit, 20% (+/- 5%)
- 20% (+/- 10%) of workplaces mentioned in scenarios and test items will be unionized.
- 10% (+/- 5%) of employers mentioned in scenarios and test items will have physical locations in more than one Canadian province.
- 10% (+/- 5%) of employers mentioned in scenarios and test items will have physical locations both inside and outside of Canada.