PUBLIC RELEASE

Technical Report: January 2022 CHRL ELE

HRPA | Human Resources Professionals Association

5 February 2022

ICKETT MEASUREMENT SYSTEMS

Contents

List of Tables

List of Figures

Executive Summary[1]

Note that this technical report covers only the primary new form or forms administered during an administration, and not detailed results for all forms used.

The CHRL Employment Law Exam (CHRL ELE) was administered to 142 candidates using computer-based testing via live remote proctoring January 19–28, 2022, 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=128[2]) was 72.6 (71.9%), and for all candidates it was 72.0 (71.3%), out of 101 validated items for scoring. Reliability was borderline at .79 (noting that there is range restriction with these candidates, and the disattenuated value was .83). The final set of scored items adhered to the blueprint parameters.

The pass mark was set using equating back to the September 2021, May 2021, and January 2021 administrations, yielding an integer pass mark of 62. 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 85.9% and a pass rate for all candidates of 85.2%.

This report, the analyses performed, and the processes followed are consistent with NCCA standards[3] and ISO 17024 standards.[4]

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

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

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

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

Administration

Form Setting

Using only validated test items, Wickett Measurement Systems prepared one 110-item test form. Wickett constructed the final test form according to the following parameters:

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

After selecting the 110 items for each form, Wickett split the forms in half to allow for the administration of the exam in two sections. Section 1 was allocated 55 items and Section 2 was allocated 55 items. With each form, the two sections were set to balance for:

- Proportion of independent items and case sets
- Number of words
- Item difficulty
- Item discrimination (adjusted point-biserial)
- Number of experimental items
- Adherence to blueprint
- Number of anchor items

The final form was reviewed for currency and enemy items by Elizabeth Austin and Nadine Bellhouse (CHRL Examination Validation Committee members), facilitated remotely, in a session held October 29, 2021.

The final form composition for the primary January 2022 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, HRPA makes use of previously validated and administered test forms along with new test forms, in addition to employing other mechanisms to maintain the integrity of the exams and candidate scores.

Table 1: Domain fit at administration

	Domain	Actual Items	Target Range	Target Items	Variance
A	Employment Contracts and Terminations	49	46% ± 5%	46–56	—
B	Employer Obligations	36	33% ± 4%	32–40	—
C	Regulations and Legislation	25	21% ± 3%	20–26	—
	TOTAL	**110**		**110**	**—**

Table 2: Cognitive level fit at administration

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	13	10% ± 3%	8–14	—
Application	53	50% ± 10%	44–66	—
Critical thinking	44	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 using live remote proctoring and at Prometric test sites primarily in Ontario. The testing window was January 19–28, 2022, inclusive, and 142 candidates wrote the exam.

Candidates were able to select either a test centre (assuming one was available reasonably close to them) or live remote proctoring from a location of their choosing. Standard security methods (as per Prometric protocols[5]) were employed for both methods. Due to lockdown measures in place, test centres were not widely available for this sitting. Candidates were allowed one 15-minute break after submitting section 1 and before beginning section 2. This break did not count against total time for the candidate.

Candidates had access to a basic-function calculator on screen 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*

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

- LRA – *Labour Relations Act, 1995*
- OHRC – *Human Rights Code*
- OHSA – *Occupational Health and Safety Act*
- PEA – *Pay Equity Act*
- WSIA – *Workplace Safety and Insurance Act, 1997*

Federal

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

The versions of the legislation were as accessed on October 21, 2021.

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 3 hours, 1 minute (4 minutes more than in September 2021). The section 1 mean time was 1 hour, 34 minutes; the section 2 mean time was 1 hour 27 minutes. Seventeen candidates (12%) took the full 3½ hours, suggesting that those candidates may have wanted more time, and 6 candidates (4%) left at least one item blank, suggesting that those candidates timed out of the exam before being able to complete it. These metrics will continue to be monitored, but presently they do not appear problematically high. Compared with the September 2021 administration, there was little change in these values. Note that because they have access to legislation, candidates may take more time than intended by researching more answers. This may generally skew time metrics higher.

Some candidates who were granted an accommodation took longer than the 3½ hours.

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

Candidate scores were computed across the window to look for any evidence of item exposure. As shown in Figure 2, there was little variation across the window, and the difference between the first 2 days and the last 2 days was a small decrease of 0.5 marks out of 110.

As a matter of interest, candidate volumes were also examined across the window; these are also shown in Figure 2. As is usually the case, candidates were more likely to book their session in the last three days of the testing window.

Figure 1: Examination time distribution for all candidates

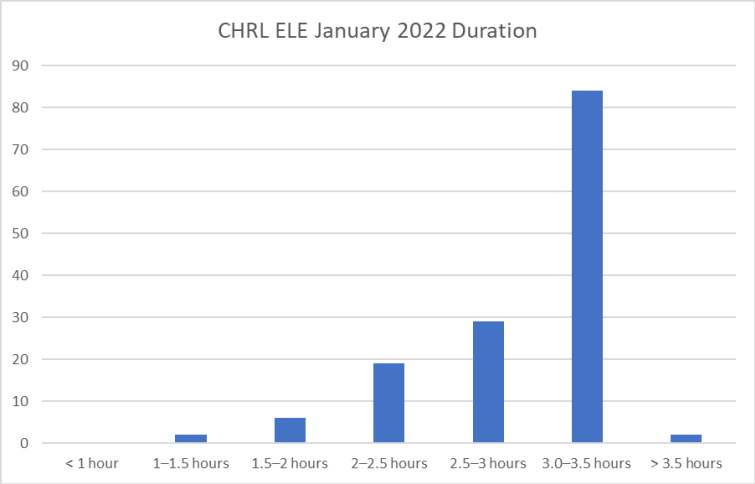
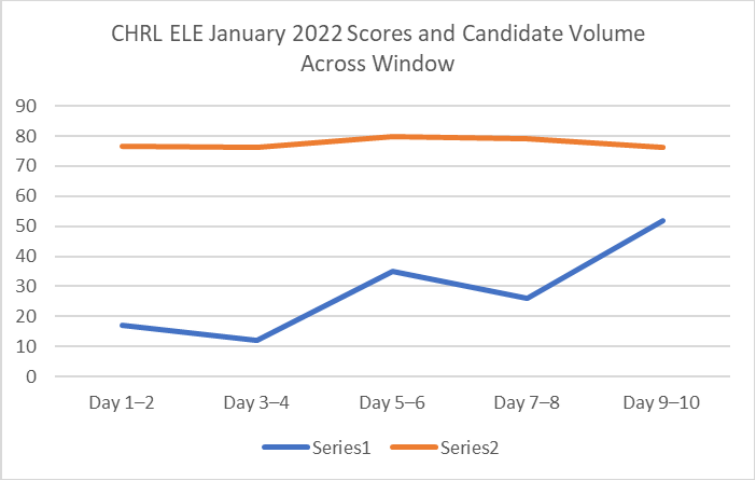


Figure 2: Candidate volume and score trends across testing window



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. No 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. All outliers were removed from initial analyses; candidates with abnormal response patterns (such as having 5 or more blanks) were also removed. 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) scored lower than did those who had not (64.9 and 71.1, respectively, on the full exam of 110 items). This difference was statistically significant ($t(18)=3.84$, $p<.01$), 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 128 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; 140 candidates responded (response rate, 98.6%).

Table 3 shows the content-related questions; there was a tendency to more neutrality on these questions though several show moderately high positive ratings. A drop in ratings for the questions being clearly written (Q4) will be monitored. The endorsement of having sufficient time dipped this administration as well, without a meaningful increase in testing time; it will continue to be monitored. Table 4 shows the responses to the administration-related questions. Note that candidates were generally positive about the administration experience, though there was drop in ratings for ease of access to legislation and case texts (Q14) that will be monitored.

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

	Question	SA	A	N	D	SD	Score	Agree rate	*Agree rate last 5^*
1.	The time allotted for this examination was sufficient.	41	54	20	17	8	3.74	**68%**	*78%*
2.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	34	77	18	11	0	3.96	**79%**	*80%*
3.	I feel I was adequately prepared to write this examination.	12	81	37	10	1	3.66	**66%**	*63%*
4.	The questions in the examination were clearly written.	12	65	37	24	3	3.42	**55%**	*60%*
5.	The terminology used in the examination was accurate.	16	94	22	9	0	3.83	**78%**	*82%*
6.	The situations presented in the examination were realistic.	33	84	16	6	1	4.01	**84%**	*86%*
7.	The questions in the examination reflected the Employment Law Examination blueprint.	16	83	29	11	1	3.73	**71%**	*69%*
8.	The examination was a fair assessment of my ability.	12	67	37	17	6	3.45	**57%**	*56%*

*Response categories: SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.
^Mean value of candidate agreement across the previous 5 administrations.

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

	Question	SA	A	N	D	SD	Score	Agree rate	*Agree rate last 5^*
9.	I was able to book to write the examination at a time that was convenient for me.	50	70	6	8	6	4.07	**86%**	*82%*
10.	I was well informed about the examination rules and regulations.	58	77	3	2	0	4.36	**96%**	*96%*
11.	Proctors enforced the exam-day rules.	75	62	3	0	1	4.49	**97%**	*97%*
12.	Proctors were professional and courteous.	80	49	7	4	1	4.44	**91%**	*95%*
13.	The tutorial helped me understand how to complete the examination on the computer.	56	68	13	4	0	4.25	**88%**	*91%*
14.	The legislation and case texts were easy to access during the examination.	29	55	12	38	7	3.43	**60%**	*66%*
15.	Navigation through the examination was easy and intuitive.	40	79	10	11	0	4.06	**85%**	*90%*

*Response categories: SA = Strongly Agree; A = Agree; N = Neutral; D = Disagree; SD = Strongly Disagree.
^Mean value of candidate agreement across the previous 5 administrations (excepting Q9 – Q11 which have only been asked in this form since August 2020).

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

Table 5: Testing location

Response	Count	%
Test centre	21	15%
Own location	119	85%

Candidates who indicated they tested in the own location (via live remote proctoring) responded to questions shown in Table 6 through Table 9. These candidates favoured using their own location versus a test centre, and they were generally positive about the experience and felt that HRPA should continue to offer the option in the future. As expected, COVID-19 related concerns were a motivating factor for many in choosing live remote proctoring, and convenience of not having to travel was also a main driver. Lockdown measures likely led to many responding that the reason they did not choose a test centre was because of very few centres being open.

Table 6: Preferred location (live remove proctoring candidates)

Response	Count	%
I preferred using my own location.	97	82%
I preferred going to a test centre.	9	8%
I have no preference.	13	11%

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

Response	Count	%
No test centres were open in my area.	26	22%
I preferred to avoid being around other people.	16	13%
I liked the convenience of not having to travel to a test centre.	54	45%
I felt like I would perform better in my own environment.	14	12%
Other (please specify)	9	8%

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

	Count	%
Very positive	33	28%
Positive	50	42%
Neutral	21	18%
Negative	14	12%
Very negative	1	1%

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

Response	Count	%
Yes	118	99%
No	1	1%

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

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

	Count	%
Strongly agree	5	24%
Agree	13	62%
Neither agree nor disagree	1	5%
Disagree	2	10%
Strongly disagree	0	0%

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

Response	Count	%
Yes	19	90%
No	2	10%

Open-ended questions were also posed to candidates asking for any additional comments in general and regarding test delivery method. Those comments were provided to HRPA for information and consideration. Nothing actionable with respect to scoring emerged in these comments.

Initial Analysis

The full 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 12. The section statistics are shown in Table 13.

Table 12: Initial examination statistics

Index	CHRL ELE
Items	110
Total candidates	142
Candidates in analysis	128
Mean	78.4 (71.3%)
Range	58–98 (52.7–89.1%)
Standard deviation	9.22
Cronbach's alpha	.78
Disattenuated alpha	.85
Mean r_{pb}*	.16

Table 13: Section item statistics

Index	Section 1	Section 2
Scorable items	55	55
Candidates in alysis	128	
Mean	38.6 (70.1%)	39.1 (71.2%)
Standard devin	5.2	5.3
Range	25–49	26–49
Mean time (minues)	94.1	87.0

A comparison between live remote proctoring candidates (mean score = 77.3) and test centre candidates (mean score = 79.0) was made which showed a non-significant difference favouring test centre candidates ($t(141)=0.80$, *ns*)

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

Standard classical test theory analysis was conducted to identify the following:

1. Item difficulty (percent obtaining correct result, *p*)
2. Item discrimination (corrected point-biserials, r_{pb}*)

3. Distractor quality (based primarily on distractor discrimination)

Wickett compiled these statistics, along with any comments made by candidates concerning flagged items, to identify items that may have been keyed incorrectly or that were performing poorly. Most emphasis was placed on the corrected point-biserials as evidence of item quality, after removing items 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 February 2, 2022, using members of the CHRL Examination Validation Committee (EVC). The group (Table 14) 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 14: CHRL Examination Validation Committee – Key validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
Nancy Richard (CHAIR)	CHRL	15–19	2017	Canadian Nuclear Safety Commission
Jennifer King (VICE-CHAIR)	CHRL	20–29	2017	Banking Industry
Liz Austin	CHRL	10-15	2021	Unionized
Nadine Bellhouse	CHRL	15–19	2019	Printing
Jennifer Borges	CHRL	10–14	2017	Manufacturing
✓ **Tanya Dacres**	CHRL	15-19	2021	Digital Business/ Transformation
Annette Dhanasar	CHRL	15–19	2017	Real Estate
✓ **Maja Falarz**	CHRL	5-9	2017	Stock Exchange
✓ **Christine Kelsey**	CHRL	5–9	2017	Media
Cynthia Ogbarmey-Tetteh	CHRL	15-19	2021	Public/Education and Government
✓ **Karen Pantaleo**	CHRL	20–29	2019	Healthcare / Consulting
✓ **Kristin Rivait**	CHRL	15–19	2017	Manufacturing
✓ **Jim St. Germain**	CHRL	20-25	2021	First Nations/ Municipal/ Manufacturing
Laurie Torno	CHRL	20–29	2018	Education – University focus on pensions/benefits/ compensation

✓Participated in the session.

The group was informed that test reliability, as measured by Cronbach's alpha, was .78 based on the set of 110 potentially scored items and that this was below the generally accepted threshold of .80. The group was advised that restriction of range was considered the most likely basis for the lower value and were provided with the disattenuated value of .85 as an estimate of the true reliability of these test scores. They were also informed that part of the goal of the key validation review was to bring this value up if possible.

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

Not all remaining items were strong-performing, and several items were retained that were very easy or very hard or that had a low corrected point-biserial. Most were moderate to strong items, however. The final alpha for the set of 101 scored items was .79 (disattenuated alpha was .83). The difficulties ranged from 36.7% to 96.1%, with a mean of 71.9%. The r_{pb}* values ranged from −.11 to .43, with a mean of .17. Note that with a small sample of candidates, negative point-biserial values are not necessarily a sign of a problematic item, and items that have performed well in the past were more likely to be retained even if showing a poor point-biserial in this candidate sample.

Table 15 shows the scored CHRL ELE's final fit to the domain weighting. Table 16 shows the same for cognitive level, and Table 17 shows the same for item type. The exam fit on all dimensions.

The group endorsed the final set of items for use in scoring the January 2022 CHRL ELE candidates.

Table 15: Domain fit for final scored items

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	41	46% ± 5%	41–52	—
B Employer Obligations	35	33% ± 4%	29–38	—
C Regulations and Legislation	25	21% ± 3%	18–25	—
TOTAL	**101**		**101**	**—**

Table 16: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	11	10% ± 3%	7–14	—
Application	48	50% ± 10%	40–61	—
Critical thinking	42	40% ± 10%	30–51	—
TOTAL	**101**		**101**	**—**

Table 17: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	23	25% ± 3%	22–29	—
Case	78	75% ± 3%	72–79	—
TOTAL	**101**		**101**	**—**

Establishing the Pass Mark: Equating

Equating, as per Kolen and Brennan (2014)[6] and Livingston and Kim (2009),[7] was used to establish the pass mark for the January 2022 CHRL ELE. The goal of this process was to set a pass mark 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 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 (January 2021, May 2021, and September 2021). The intention following these equating runs was to average them to arrive at a final pass mark for the January 2022 CHRL ELE.

Equating Back to the January 2021 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. Because both January administrations did have close to 100 candidates, all results were considered in this analysis.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta-plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 *SD* units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form that adhered to the blueprint. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.71 and a mean corrected point-biserial of .19.

Table 18 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.

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

Table 18: Anchor item fit to blueprint – To January 2021

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

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker was considered the primary method, but there were also arguments for Levine observed-score method because of differences in the anchor means. Further, because the administrations had just over 100 candidates, the mean and circle arc methods were also considered.

Table 19 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 January 2022 CHRL ELE scored lower than the sample taking the January 2021 CHRL ELE (71.0% vs. 74.2%, respectively; $t(250)=2.37$, $p<.05$). Because the January 2022 CHRL ELE candidates were of lower ability (based on the anchors), they should have a lower pass rate (however, the tails of the distribution will be more erratic with small samples).

The equating analysis shows this result (Table 20). All methods show a pass mark of 60–64, with the large fluctuations due to substantial variance differences (likely partly attributable to the small number of candidates). Given the sample sizes involved, Tucker or Levine observed would be the primary methods under consideration though a case could be made for the mean and circle arc methods because of the small samples. Tucker was provisionally selected from this analysis based on precedent.

Table 19: Equating parameter table – To January 2021

		Jan. 2021	Jan. 2022
	N	124	128
	Scored items	101	101
Mean score	Total	74.4%	71.9%
Mean score	Anchors	74.2%	71.0%

Table 20: Equating outcome table – To January 2021

	Pas Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equatingn. 2021	**63.37**	**64**	**91.1%**	**95.2%**
Tucker	59.61	60	*91.5%*	*90.6%*
Levine oserved	60.75	61	*88.0%*	*87.5%*
Mean	62.88	63	*82.4%*	*82.8%*
Circle Ar1	63.50	64	*81.0%*	*82.0%*
Circle Ar2	63.50	64	*81.0%*	*82.0%*

Equating Back to the May 2021 Administration

Linear equating (Tucker) was the chosen method for setting the pass mark and it was conducted once key validation was complete. Linear equating is the primary method considered with more than 100 candidates; equipercentile equating would have been considered with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc equating is most prudent.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta-plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 *SD* units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form that adhered to the blueprint. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.72 and a mean corrected point-biserial of .18.

Table 21 shows the fit of the set of anchor items to the blueprint, as percentages. The actual counts are reasonably aligned with the targets and reflect the scope and approximate weighting across the full exam.

Table 21: Anchor item fit to blueprint – To May 2021

	Area	Actual	Target
A	Employment Contracts and Terminations	48%	46%
B	Employer Obligations	28%	33%
C	Regulations and Legislation	24%	21%

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker was considered the primary method (though with consideration being paid also to the mean and circle arc methods because of the small number of candidates in January 2022).

Table 22 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 January 2022 CHRL ELE scored about the same as the sample taking the May 2021 CHRL ELE (71.5% vs. 71.8%, respectively; $t(355)=0.22$, *ns*). Because the January 2022 CHRL ELE candidates were of slightly lower ability (based on the anchors), they should have a slightly lower pass rate.

The equating analysis showed this result for the most part, though with a small dip in pass rate that could be attributed to more pronounced tails in the scoring distribution for January 2022 candidates (Table 23). The methods showed an integer pass mark of 61–63. Given the sample sizes involved, Tucker would be the primary method under consideration and the equated value of 60.5 was carried forward in the analysis.

Table 22: Equating parameter table – To May 2021

		May 2021	Jan. 2022
	N	229	128
	Scored items	102	101
Mean score	Total	72.6%	71.9%
Mean score	Anchors	71.8%	71.5%

Table 23: Equating outcome table – To May 2021

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equating May 2021	**63.54**	**64**	**91.5%**	**91.7%**
Tucker	60.46	61	*88.0%*	*87.5%*
Levine observed	61.08	62	*85.2%*	*85.9%*
Mean	62.29	63	*82.4%*	*82.8%*
Circle Arc 1	62.49	63	*82.4%*	*82.8%*
Circle Arc 2	62.49	63	*82.4%*	*82.8%*

Equating Back to the September 2021 Administration

Linear equating (Tucker) was the chosen method for setting the pass mark and it was conducted once key validation was complete. Linear equating is the primary method considered with more than 100 candidates; equipercentile equating would have been considered with more than 1,000 candidates. With candidate samples of fewer than 100, mean or circle arc equating is most prudent.

All candidates in the analysis (i.e., no repeat candidates or outliers) were used in the equating process. Delta-plot analysis was used to identify anchor items showing substantial deviations (generally, although not exclusively, greater than 3 *SD* units) from expected difficulty values, with an emphasis on establishing an anchor set with difficulty equivalent to that of the full form that adhered to the blueprint. Further, items with very high or low difficulty values and those with low corrected point-biserials were also flagged for potential removal from the anchor set. The goal was a strong midi-test (i.e., moderate range of difficulty, moderate to high discrimination, fit to blueprint) of sufficient length to estimate candidate ability.

The selected set of anchor items had a mean difficulty of 0.71 and a mean corrected point-biserial of .21.

Table 24 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 24: Anchor item fit to blueprint – To September 2021

	Area	Actual	Target
A	Employment Contracts and Terminations	44%	46%
B	Employer Obligations	32%	33%
C	Regulations and Legislation	24%	21%

The mean, Tucker, Levine observed-score, and circle arc methods were computed to ascertain concordance of solutions. Given the sample sizes and similarities of test parameters, Tucker was considered the primary method (though with consideration being paid also to the mean and circle arc methods because of the small number of candidates in January 2022).

Table 25 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 January 2022 CHRL ELE scored essentially the same as the sample taking the September 2021 CHRL ELE (71.2% vs. 71.6%, respectively; $t(330)=0.28$, *ns*). Because the January 2022 CHRL ELE candidates were of about the same ability (based on the anchors), they should have approximately the same pass rate (though the tails of the distribution will be more erratic with small samples).

The equating analysis shows this result (Table 26). All methods show a pass mark of 62–63. Given the sample sizes and comparability of anchor parameters, Tucker would be the primary methods under consideration.

Table 25: Equating parameter table – To September 2021

		Sep. 2021	Jan. 2022
	N	204	128
	Scored items	102	101
Mean score	Total	73.9%	71.9%
	Anchors	71.6%	71.2%

Table 26: Equating outcome table – To September 2021

	Pas Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equatin Sep. 2021	**65.27**	**66**	**85.8%**	**89.2%**
Tucker	61.89	62	*85.2%*	*85.9%*
Levine oserved	62.32	63	*82.4%*	*82.8%*
Mean	62.75	63	*82.4%*	*82.8%*
Circle Ar1	62.88	63	*82.4%*	*82.8%*
Circle Ar2	62.85	63	*82.4%*	*82.8%*

Combined Results

Table 27 shows the pass mark values across the equating runs. The value highlighted in green is the one that would be selected based on sample parameters at each equating run. For the September and May 2021 analyses, the Tucker value was the most suitable choice, but for the January 2021 analysis any of three values could be supported. Further, because of the low number of candidates in January 2022, the mean or circle arc methods are a viable choice in all of the analyses.

In the end, the goal is to identify the best representation of what the pass mark should be based on the available information. In this situation, the weighted average of the Tucker values is showing as 2 points below the mean and circle arc methods, and 1 point below the Levine observed-score method. Given the relatively small sample in January 2022, a case can be made for adding weight to the mean and circle arc methods. Further, there was a great disparity in variances across administrations, which served to push the Tucker estimates lower than may be attributable to the exam itself. To strike a balance between the Tucker results and mean/circle arc methods, the simple average of 60.7 from the Tucker weighted average and the 62.9 from the circle arc 1 weighted average was calculated at 61.8 and became the recommended pass mark. This approach serves to reduce the heavy influence of variance differences between administrations.

Using the established convention for this testing program, the averaged pass mark would be rounded up to a cut score of 62. The resulting pass rate for first-time candidates (85.9%) is about the same as in recent administrations (though marginally lower), which is in line with expectations from the equating runs. The pass rate for all candidates was 85.2%. See Table 28 for historical pass rates.

The final pass mark value, and the process used to derive it, was presented to the CHRL EVC (Table 29) via teleconference on February 4, 2022. No concerns were raised regarding the process, pass mark or pass rate. The panel formally approved the pass mark (which was presented along with the consequent pass rate data) for recommendation to HRPA. The HRPA

Registrar accepted the recommended pass mark and so the pass mark was formally established.

Table 27: Equating outcome table – Combined results

	Jan. 21	May 21	Sep 21	Weighted Average
Tucker	59.6	60.5	61.9	60.7
Levine observed	60.8	61.1	62.3	61.4
Mean	62.9	62.3	62.7	62.6
Circle arc 1	63.5	62.5	62.9	62.9
Circle arc 2	63.5	62.5	62.8	62.9

Table 28: Historical pass rates

	Pass rate	
	All	**First-time**
Jan. 2019	85.7%	85.9%
May	83.5%	86.0%
Sep.	87.5%	89.3%
Jan. 2020	86.8%	89.6%
Aug.	88.2%	90.7%
Nov.	83.2%	84.7%
Jan. 2021	91.1%	95.2%
May	91.5%	91.7%
Sep.	85.8%	89.2%
Jan. 2022	85.2%	85.9%

Table 29: CHRL Examination Validation Committee – Pass mark approval

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
Nancy Richard (CHAIR)	CHRL	15–19	2017	Canadian Nuclear Safety Commission
✓ **Jennifer King (VICE-CHAIR)**	CHRL	20–29	2017	Banking Industry
✓ **Liz Austin**	CHRL	10-15	2021	Unionized
✓ **Nadine Bellhouse**	CHRL	15–19	2019	Printing
✓ **Jennifer Borges**	CHRL	10–14	2017	Manufacturing
Tanya Dacres	CHRL	15-19	2021	Digital Business/ Transformation
Annette Dhanasar	CHRL	15–19	2017	Real Estate
Maja Falarz	CHRL	5-9	2017	Stock Exchange
Christine Kelsey	CHRL	5–9	2017	Media
Cynthia Ogbarmey-Tetteh	CHRL	15-19	2021	Public/Education and Government
Karen Pantaleo	CHRL	20–29	2019	Healthcare / Consulting
Kristin Rivait	CHRL	15–19	2017	Manufacturing
✓ **Jim St. Germain**	CHRL	20-25	2021	First Nations/ Municipal/ Manufacturing
✓ **Laurie Torno**	CHRL	20–29	2018	Education – University focus on pensions/benefits/ compensation

✓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 30 provides the means and standard deviations for the domains and for the total score, using all candidates who took the January 2022 CHRL ELE. Table 31 provides the correlations between each domain. Figure 3 shows the distribution of scores for all candidates, along with the pass mark.

Table 30: Total and domain scores for all candidates

	Domain	Percentage	Mean	SD*
A	Employment Contracts and Terminations	73%	29.8	4.7
B	Employer Obligations	70%	24.5	3.5
C	Regulations and Legislation	71%	17.7	2.6
	Total score	71.3%	72.0	9.0

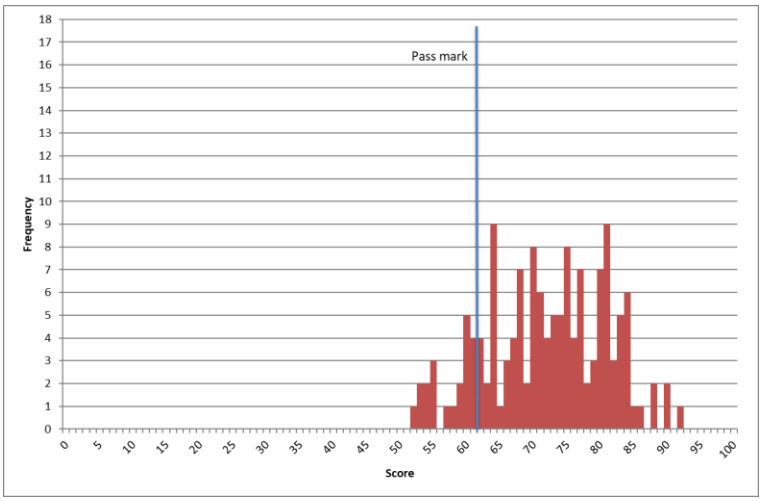
*SD = Standard deviation.

Table 31: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.54	.56
B			.48
C			

*See Table 30 for the full name of each functional area.

Figure 3: Score distribution for all candidates



Key Examination Metrics

Table 32 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 32: Key examination metrics – Candidates included in analysis only

Index	January 2022	September 2021	May 2021	January 2021	November 2020
Scored items	101	102	102	101	101
Candidates	128	204	229	124	255
Mean	72.6 (71.9%)	75.4 (73.9%)	74.0 (72.6%)	75.20 (74.4%)	73.0 (72.3%)
Median	74 (73.3%)	75 (73.5%)	75 (73.5%)	76 (75.2%)	74 (73.3%)
Skewness	−0.322	−0.107	−0.520	−0.387	−0.328
Kurtosis	−0.561	−0.084	0.343	0.140	−0.439
Range	52–92 (51.5–91.1%)	54–95 (52.9–93.1%)	52–91 (51.0–89.2%)	55–90 (54.5–89.1%)	50–92 (49.5–91.1%)
Standard deviation	9.15	7.98	7.61	6.81	8.82
Cronbach's alpha	.79	.74	.71	.66	.79
Mean r_{pb}*	.17	.14	.13	.12	.17
SEM[i]	4.16	4.06	4.13	3.97	4.06
SEM at the pass mark	4.60	4.51	4.55	4.45	4.47
Decision consistency (uncorrected)[ii]	.89	.87	.90	.91	.87
Perceived fairness[iii]	57%	55%	50%	54%	53%
Pass mark	61.832	65.270	63.543	63.373	63.373
Effective pass mark	62	66	64	64	64
Pass rate	85.9%	89.2%	91.7%	95.2%	84.7%

[i]SEM = standard error of measurement.

[ii]Subkoviac method.

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

Related Development Activities

Since the last administration of the CHRL ELE in September 2021, the following exam development activities have taken place.

Validation

To renew the validation of items expiring from usability and to validate newly written items, a validation session was held with the EVC (see Table 33) remotely on October 15, 2021. This session was an extension of the previous validation session held in September.

Table 33: CHRL Examination Validation Committee – Validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓ **Liz Austin**	CHRL	10-15	2021	Unionized
✓ **Nadine Bellhouse**	CHRL	15–19	2019	Printing
✓ **Jennifer Borges**	CHRL	10–14	2017	Manufacturing
✓ **Tanya Dacres**	CHRL	15-19	2021	Digital Business/ Transformation
Annette Dhanasar	CHRL	15–19	2017	Real Estate
Maja Falarz	CHRL	5-9	2017	Stock Exchange
Christine Kelsey	CHRL	5–9	2017	Media
✓ **Jennifer King**	CHRL	20–29	2017	Banking Industry
✓ **Cynthia Ogbarmey-Tetteh**	CHRL	15-19	2021	Public/Education and Government
✓ **Karen Pantaleo**	CHRL	20–29	2019	Healthcare / Consulting
Nancy Richard	CHRL	15–19	2017	Canadian Nuclear Safety Commission
Kristin Rivait	CHRL	15–19	2017	Manufacturing
Jim St. Germain	CHRL	20-25	2021	First Nations/ Municipal/ Manufacturing
✓ **Laurie Torno**	CHRL	20–29	2018	Education – University focus on pensions/benefits/ compensation

✓Participated in the session.

The EVC members received advance materials outlining:

- Purpose of the session

- Description of the CHRL credential
- CHRL ELE blueprint
- Criteria for good test items
- Validation process
- Relevant legislation

The committee members received refresh training on the validation activity on the first day of the session. Committee members were provided with 44 items and case texts via a secure file share site, and then worked individually reviewing items through the day, submitting their appraisal and any suggested revisions to Wickett through the day. They were directed to make sure the items reflected current practice and were suitable to make decisions about who should receive the CHRL credential.

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

For each item, the committee was asked to either:

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

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

The committee validated 43 items and case texts as suitable for the CHRL ELE, rejected 1 item, and shifted 0 items for eligibility in the CHRP ELE bank. Fifteen items/case texts were revised prior to validation as part of this exercise. The committee also verified the topic and cognitive level for all items, and added rationales and references where missing, incomplete, or not current.

Appendix

Blueprint

CHRL Employment Law Examination Blueprint

Human Resources Professionals Association

Version 2.0
Approved by CHRL Exam Validation Committee March 13, 2018
Approved by HRPA 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 34.

Table 34: 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 35.

Table 35: Employment Law Examination Blueprint Content Weights

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

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

Note: Reasonable ranges around the Topic weights are employed.

Cognitive Level

The cognitive level weights are based on Bloom's taxonomy. The purpose of this weighting is generally to ensure that an examination does not unintentionally over-focus on specific types of items, and to provide candidates with a range of items (in approximate proportion) that reflects the cognitive operations they must apply on the job. The weights appear in Table 36.

Table 36: 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.

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.