PUBLIC RELEASE

Technical Report: August 2020 CHRL ELE

Human Resources Professionals Association

19 October 2020

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 (which may include previously used forms, scrambled forms, and other modifications to maintain exam and score integrity).

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

The CHRL Employment Law Exam (CHRL ELE) was administered to 169 candidates using computer-based testing via live remote proctoring and at Prometric test centres August 17–23, 2020, 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=151[2]) was 75.5 (74.8%), and for all candidates it was 74.8 (74.0%), out of 101 validated items for scoring. Reliability was borderline at .74 (noting that there is substantial range restriction with these candidates, and the disattenuated value was .84). The final set of scored items adhered to the blueprint parameters.

The pass mark was set using equating back to the May 2019, September 2019 and January 2020 administrations, yielding an integer pass mark of 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 90.7% and a pass rate for all candidates of 88.2%.

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

The final form was reviewed for currency and enemy items by Jennifer Borges and Kristin Rivait (CHRL Examination Validation Committee members), facilitated remotely, in a session held February 25, 2020.

The final form composition for the primary August 2020 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	50	46% ± 5%	46–56	—
B	Employer Obligations	38	33% ± 4%	32–40	—
C	Regulations and Legislation	22	21% ± 3%	20–26	—
	TOTAL	**110**		**110**	**—**

Table 2: Cognitive level fit at administration

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	10	10% ± 3%	8–14	—
Application	52	50% ± 10%	44–66	—
Critical thinking	48	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 August 17–23, 2020, inclusive, and 169 candidates wrote the exam.[5]

Candidates were able to select either a test centre (assuming one was available reasonably close to them) or live remote proctoring from a location of their choosing. Standard security methods (as per Prometric protocols[6]) were employed for both methods. Candidates were allowed up to one unscheduled10-minute break during the examination (the examination timer was not stopped during this break).

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

Provincial

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

[5] Due to technical difficulties requiring the rescheduling of some candidates, testing continued on August 24.

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

Federal

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

The versions of the legislation were as accessed on February 14, 2020.

Analysis

Data Cleaning and Integrity Checks

Prometric provided data in .xml format via a secure ftp site. Candidate files were provided as candidates completed the examination throughout the testing window. These files were extracted to Microsoft Excel for processing. They contained identifying information for each candidate, form information, start and stop times, answer string, key string, candidate total score, item comments if the candidate made any, and time spent per item.

The data files received were reconciled against the roster provided by Prometric to ensure that all .xml files had been received. Further, each candidate total score as computed by Prometric was reconciled with that computed by Wickett for the full set of 110 items to verify key accuracy. Comments on items were also reviewed to identify any specific item-level issues. No problems were encountered.

The average time taken by all candidates was assessed to detect potential examination timing concerns. The distribution is shown in Figure 1. The mean was 2 hours, 58 minutes (7 minutes less than in January 2020). Note that minor variances in how time is recorded have been seen over the past few administrations, and so it is unclear if time differences are reflecting a technical issue or an actual change in candidate behaviour relative to the examination.

Eighteen candidates (11%) took the full 3½ hours, suggesting that those candidates may have wanted more time. Six 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 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 essential zero 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 negligible increase of 0.2 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 this was a shorter administration window that also included live remote proctoring, the normal pattern of increasing volume at the end of the window was not expected nor observed.

Figure 1: Examination time distribution for all candidates

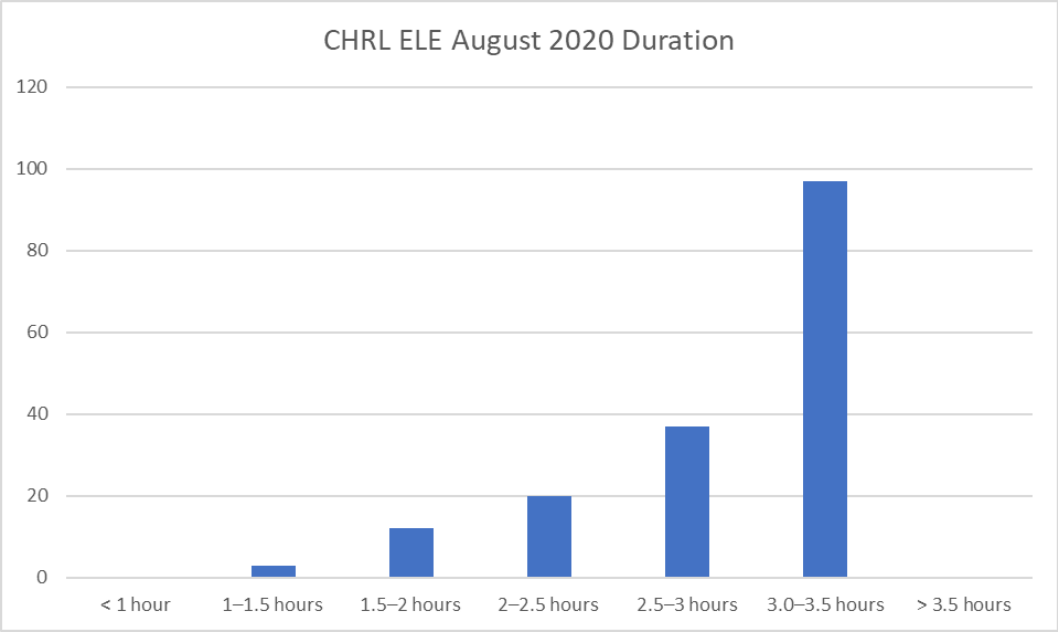
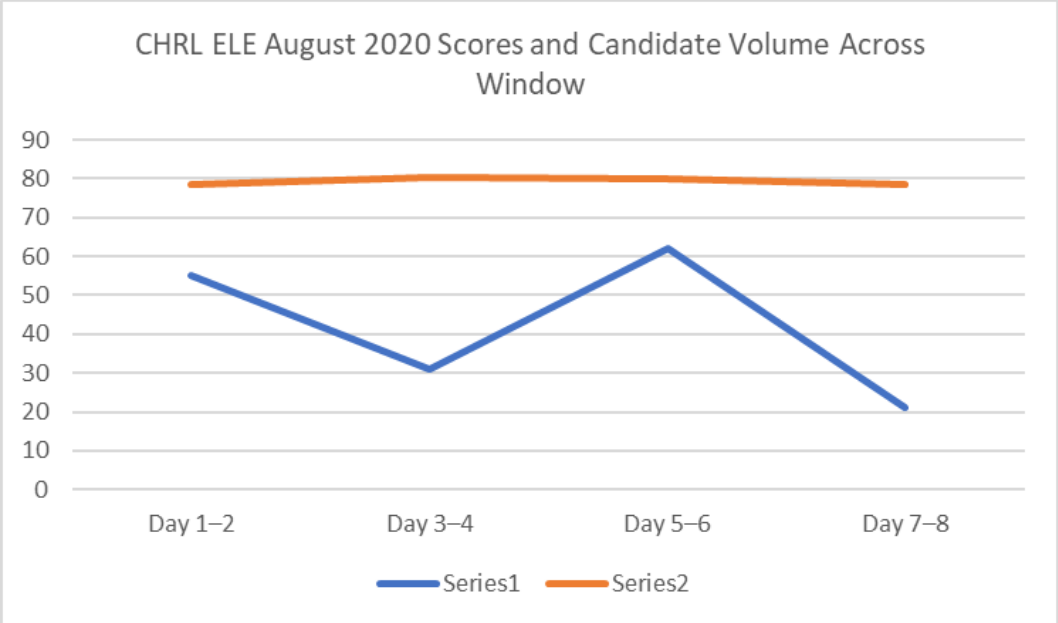


Figure 2: Candidate volume and score trends across testing window



After removing candidates who were administered a previously used test form (who were scored using the same decisions employed at the time that form was originally used), scores were calculated for all remaining candidates based on the full set of 110 items. One candidate was flagged for an abnormally low or high score (*z* value outside +/− 3.0). Also, the 110 items were arbitrarily broken into 4 blocks of 25 items for each candidate plus 1 final block of 10 items; the 5 resulting subscores for each candidate were evaluated for outliers as well. For candidates

with any subscore more than 3 standard deviations (SD) from their average *z*-score, the .xml file was examined closely for any issues. 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, 4 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 (74.5 and 79.9, respectively, on the full exam of 110 items). This difference was statistically significant ($t(19)=2.59$, $p<.05$), 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 151 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; 167 candidates responded (response rate, 99%). The survey was revised for the August administration to collect information specifically applicable to the use of live remote proctoring.

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

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

	Question	SA	A	N	D	SD	Score	Agreement
1.	The time allotted for this examination was sufficient.	64	66	10	19	8	3.95	**78%**
2.	Information available prior to exam day provided me with adequate details about the content and format of the exam.	42	87	24	14	0	3.94	**77%**
3.	I feel I was adequately prepared to write this examination.	19	104	32	11	1	3.77	**74%**
4.	The questions in the examination were clearly written.	10	100	36	20	1	3.59	**66%**
5.	The terminology used in the examination was accurate.	22	118	24	3	0	3.95	**84%**
6.	The situations presented in the examination were realistic.	34	115	13	5	0	4.07	**89%**
7.	The questions in the examination reflected the Employment Law Examination blueprint.	24	98	35	9	0	3.83	**73%**
8.	The examination was a fair assessment of my ability.	13	96	42	15	1	3.63	**65%**

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

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

	Question	SA	A	N	D	SD	Score	Agreement
9.	I was able to book to write the examination at a time that was convenient for me.	65	87	5	6	4	4.22	**91%**
10.	I was well informed about the examination rules and regulations.	70	88	5	3	1	4.34	**95%**
11.	Proctors enforced the exam-day rules.	93	68	5	1	0	4.51	**96%**
12.	Proctors were professional and courteous.	105	54	6	1	1	4.56	**95%**
13.	The tutorial helped me understand how to complete the examination on the computer.	81	70	11	3	2	4.35	**90%**
14.	The legislation and case texts were easy to access during the examination.	45	71	18	27	6	3.73	**69%**
15.	Navigation through the examination was easy and intuitive.	58	99	7	2	1	4.26	**94%**

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

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

Table 5: Testing location

Response	Count	%
Test centre	28	17%
Own location	138	83%

Candidates who indicated they tested in the own location (via live remote proctoring) responded to questions shown in Table 6 through Table 9. These candidates were largely split on whether they preferred using their own location versus a test centre, though they were generally positive about the experience and felt that HRPA should continue to offer the option in the future. As expected, COVID-19 related concerns were a motivating factor for many in choosing live remote proctoring (shown primarily in the 'other' write in responses), and lack of open test centres was also a main driver.

Table 6: Preferred location (live remove proctoring candidates)

Response	Count	%
I preferred using my own location.	66	48%
I preferred going to a test centre.	48	35%
I have no preference.	24	17%

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

Response	Count	%
No test centres were open in my area.	42	30%
I preferred to avoid being around other people.	33	24%
I liked the convenience of not having to travel to a test centre.	39	28%
I felt like I would perform better in my own environment.	4	3%
Other (please specify)	20	14%

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

	Count	%
Very positive	33	24%
Positive	50	36%
Neutral	31	22%
Negative	17	12%
Very negative	7	5%

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

Response	Count	%
Yes	132	96%
No	6	4%

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

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

	Count	%
Strongly agree	8	29%
Agree	11	39%
Neither agree nor disagree	2	7%
Disagree	5	18%
Strongly disagree	2	7%

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

Response	Count	%
Yes	20	71%
No	8	29%

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.

Table 12: Initial examination statistics

Index	CHRL ELE
Items	110
Total candidates	169
Candidates in analysis	151
Mean	80.2 (72.9%)
Range	54–95 (49.1–86.4%)
Standard deviation	7.76
Cronbach's alpha	.71
Disattenuated alpha	.86
Mean r_{pb}*	.13

A simple comparison between scores obtained by test centre candidates (mean score of 69.8%) and live remote proctoring candidates (mean score of 72.6%) was made to evaluate if there was any problematic difference in performance. The small number of candidates means this analysis is inconclusive, but a lack of significant effect ($t(167)=1.83$, *ns*) is at least supportive of there being no overall difference in candidate performance. There is a leaning towards higher scores from live remote proctoring candidates that will be monitored across future administrations.

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

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

1. Item difficulty (percent obtaining correct result, *p*)
2. Item discrimination (corrected point-biserials, r_{pb}*)
3. Distractor quality (based primarily on distractor discrimination)

Wickett compiled these statistics, along with any comments made by candidates concerning flagged items, to identify items that may have been keyed incorrectly or that were performing poorly. Most emphasis was placed on the corrected point-biserials as evidence of item quality, after removing items 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 August 28, 2020, using members of the CHRL Examination Validation Committee (EVC). This session was originally planned for August 26 but rescheduled to achieve quorum. The group (Table 13) was first reminded of the methods used for key validation and was oriented to the main statistics used to evaluate the quality of the CHRL ELE.

Table 13: CHRL Examination Validation Committee – Key validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓**Nadine Bellhouse**	CHRL	15–19	2019	Printing
✓**Jennifer Borges**	CHRL	10–14	2017	Manufacturing
✓**Annette Dhanasar**	CHRL	15–19	2017	Technology
✓**Debbie Hynes**	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
Karen Pantaleo	CHRL	20–29	2019	Health care and consulting
✓**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 .71 based on the set of 110 potentially scored items and that this was below the generally accepted threshold of .80. The group was advised that restriction of range was considered the most likely basis for the lower value and were provided with the disattenuated value of .86 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. 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 9 items

that they felt were least appropriate to retain for scoring. Panel members' comments about specific items were recorded for future item revision activities.

Not all remaining items were strong-performing, and several items were retained that were very easy or very hard or that had a low corrected point-biserial. Most were moderate to strong items, however. The final alpha for the set of 101 scored items was .74 (disattenuated alpha was .84). The difficulties ranged from 31.1% to 98.0%, with a mean of 74.8%. The r_{pb}* values ranged from −.17 to .40, with a mean of .14. Note that with a small sample of candidates, negative point-biserial values are not necessarily a sign of a problematic item, and items that have performed well in the past were more likely to be retained even if showing a poor point-biserial in this candidate sample.

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

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

Table 14: Domain fit for final scored items

Domain	Actual Items	Target Range	Target Items	Variance
A Employment Contracts and Terminations	45	46% ± 5%	42–51	—
B Employer Obligations	35	33% ± 4%	30–37	—
C Regulations and Legislation	21	21% ± 3%	19–24	—
TOTAL	**101**		**101**	**—**

Table 15: Cognitive level fit for final scored items

Cognitive Level	Actual Items	Target Range	Target Items	Variance
Knowledge	9	10% ± 3%	8–13	—
Application	50	50% ± 10%	41–60	—
Critical thinking	42	40% ± 10%	31–50	—
TOTAL	**101**		**101**	**—**

Table 16: Item type fit for final scored items

Item Type	Actual Items	Target Range	Target Items	Variance
Independent	25	25% ± 3%	23–28	—
Case	76	75% ± 3%	73–78	—
TOTAL	**101**		**101**	**—**

Establishing the Pass Mark: Equating

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

Equating Back to the January 2020 Administration

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

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

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 .18.

Table 17 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 17: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	44%	46%
B	Employer Obligations	34%	33%
C	Regulations and Legislation	22%	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 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 August 2020 CHRL ELE scored better than the sample taking the January 2020 CHRL ELE (73.1% vs. 71.6%, respectively; $t(274)=1.14$, *ns*). Because the August 2020 CHRL ELE candidates were of modestly higher ability (non-significance notwithstanding), they should have a modestly higher pass rate.

The equating analysis bears this out (Table 19). All methods show a pass mark of 64 or 65, with the recommended solution showing a 64 which leads to a pass rate that is modestly higher than seen in January 2020.

Table 18: Equating parameter table

		Jan. 2020	Aug. 2020
	N	125	151
	Scored items	100	101
Mean score	Total	75.0%	74.8%
Mean score	Anchors	71.6%	73.1%

Table 19: Equating outcome table

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equating Jan. 2020	**65.12**	**66**	**86.9%**	**89.6%**
Tucker	63.95	64	*89.9%*	*92.7%*
Levine observed	63.68	64	*89.9%*	*92.7%*
Mean	64.79	65	*88.2%*	*90.7%*
Circle Arc 1	64.51	65	*88.2%*	*90.7%*
Circle Arc 2	64.48	65	*88.2%*	*90.7%*

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 .17.

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

Table 20: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	48%	46%
B	Employer Obligations	33%	33%
C	Regulations and Legislation	19%	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 21 shows some of the parameters used to derive the equating estimates, along with other parameters describing the test forms. Of note is that on the anchor items, the sample taking the August 2020 CHRL ELE scored marginally higher than the sample taking the May 2019 CHRL ELE (74.3% vs. 73.5%, respectively; $t(327)=0.79$, *ns*). Because the August 2020 CHRL ELE candidates were of marginally higher ability (non-significance notwithstanding), they should have a marginally higher pass rate.

The equating analysis bears this out in general (Table 22). All methods indicate a pass mark of 65–66, with the recommended Tucker method showing a 66 with a modest increase in the pass rate from 86.0% to 90.1%.

Table 21: Equating parameter table

		May 2019	Aug. 2020
	N	178	151
	Scored items	98	101
Mean score	Total	74.6%	74.8%
Mean score	Anchors	73.5%	74.3%

Table 22: Equating outcome table

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equating May 2019	**63.38**	**64**	**83.5%**	**86.0%**
Tucker	65.85	66	*87.6%*	*90.1%*
Levine observed	65.79	66	*87.6%*	*90.1%*
Mean	65.15	66	*87.6%*	*90.1%*
Circle Arc 1	64.78	65	*88.2%*	*90.7%*
Circle Arc 2	64.77	65	*88.2%*	*90.7%*

Equating Back to the September 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 0.76 and a mean corrected point-biserial of .16.

Table 23 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 23: Anchor item fit to blueprint

	Area	Actual	Target
A	Employment Contracts and Terminations	46%	46%
B	Employer Obligations	33%	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, Tucker was considered the primary method.

Table 24 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 August 2020 CHRL ELE scored higher than the sample taking the September 2019 CHRL ELE (75.9% vs. 73.8%, respectively; $t(318)=1.94$, *ns*). Because the August 2020 CHRL ELE candidates were of higher ability (non-significance notwithstanding), they should have a higher pass rate.

The equating analysis bears this out (Table 25). 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 89.3% to 90.7%.

Table 24: Equating parameter table

		Sep. 2019	Aug. 2020
	N	169	151
	Scored items	102	101
Mean score	Total	73.5%	74.8%
Mean score	Anchors	73.8%	75.9%

Table 25: Equating outcome table

	Pass Mark		Pass Rate	
Method	**Precise**	**Integer**	**All**	**First Time**
Equating Sep. 2019	**64.97**	**65**	**87.5%**	**89.3%**
Tucker	64.09	65	*88.2%*	*90.7%*
Levine observed	63.33	64	*89.9%*	*92.7%*
Mean	63.97	64	*89.9%*	*92.7%*
Circle Arc 1	63.87	64	*89.9%*	*92.7%*
Circle Arc 2	63.87	64	*89.9%*	*92.7%*

Combined Results

Table 26 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 64.712 out of 101).[9] 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 (90.7%) is modestly higher than seen on recent previous administrations, but not substantially so. The pass rate for all candidates (88.2%) was also modestly higher than seen in recent administrations. See Table 27 for historical pass rates.

The final pass mark value, and the process used to derive it, was presented to the CHRL EVC (Table 28) via teleconference on September 3, 2020. 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 HRPA. The HRPA Registrar later approved the panel's recommendation.

[9] The equating procedure back to May 2019 yielded a pass mark almost 2 points higher than did the procedures back to September 2019 and January 2020. Though this was within expected fluctuations, it was reverified to ensure accuracy of results. Further, the rounded value of just the 2 Tucker values from September 2019 and January 2020 would still yield a value that rounds up to 65.

Table 26: Equating outcome table – Combined results

	May 19	Sep. 19	Jan. 20
Tucker	65.8	64.1	64.0
Levine observed	65.8	63.3	63.7
Mean	65.1	64.0	64.8
Circle arc 1	64.8	63.9	64.5
Circle arc 2	64.8	63.9	64.5

Table 27: 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%
Jan. 2020	86.8%	89.6%
Aug.	**88.2%**	**90.7%**

Table 28: CHRL Examination Validation Committee – Pass mark approval

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
Nadine Bellhouse	CHRL	15–19	2019	Printing
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
✓**Karen Pantaleo**	CHRL	20–29	2019	Health care and consulting
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 HRPA.

Table 29 provides the means and standard deviations for the domains and for the total score, using all candidates who took the August 2020 CHRL ELE. Table 30 provides the correlations between each domain. Caution should be exercised in interpreting differences between correlations. Variation can be explained largely by the number of items making up each domain score. Figure 3 shows the distribution of scores for all candidates, along with the pass mark.

Table 29: Total and domain scores for all candidates

	Domain	Percentage	Mean	SD*
A	Employment Contracts and Terminations	76%	34.1	4.5
B	Employer Obligations	75%	26.2	3.4
C	Regulations and Legislation	69%	14.5	2.6
	Total score	74.0%	74.8	8.2

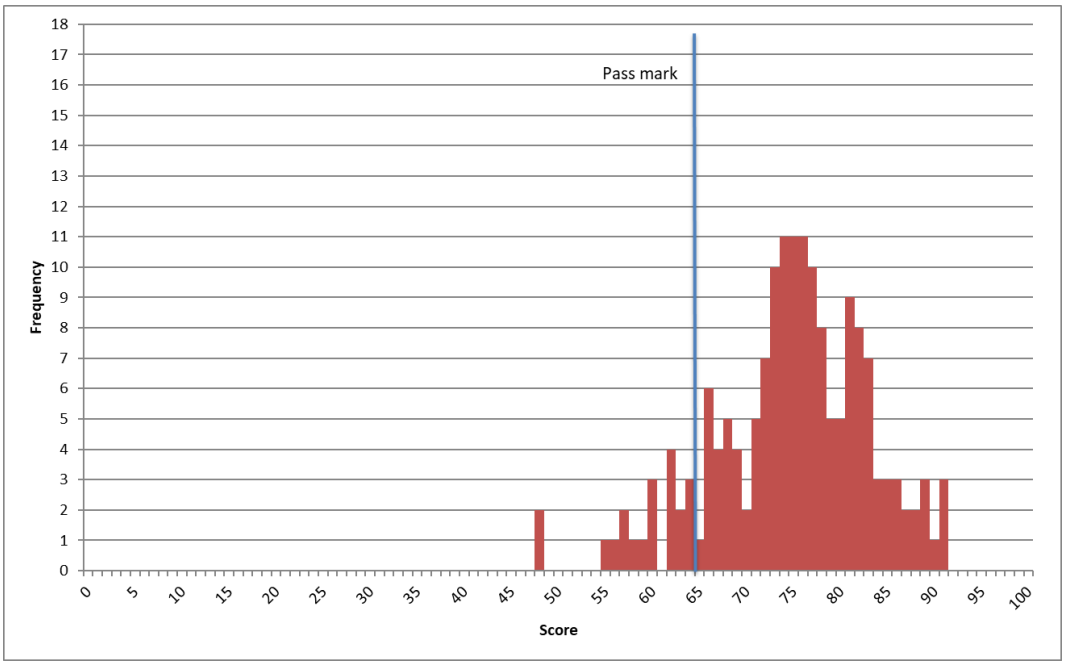
*SD = Standard deviation.

Table 30: Correlations between functional area scores for all candidates

Domain*	A	B	C
A		.47	.40
B			.39
C			

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

Figure 3: Score distribution for all candidates



Key Examination Metrics

Table 31 shows the key examination metrics for candidates included in the main analysis; that is, only first-time candidates, with outliers removed. Past metrics are provided for reference.

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

Index	August 2020	January 2020	September 2019	May 2019	January 2019
Scored items	101	100	102	98	98
Candidates	151	125	169	178	177
Mean	75.5 (74.8%)	75.0 (75.0%)	75.0 (73.5%)	73.1 (74.6%)	71.7 (73.1%)
Median	76 (75.2%)	75 (75.0%)	76 (74.5%)	74 (75.5%)	73 (74.5%)
Skewness	−0.485	−0.261	−0.595	−0.421	−0.469
Kurtosis	0.522	0.274	0.486	0.258	−0.212
Range	48–91 (47.5–90.1%)	53–94 (53.0–94.0%)	50–92 (49.0–90.2%)	47–91 (48.0–92.9%)	46–90 (46.9–91.8%)
Standard deviation	7.75	7.12	8.11	8.20	9.20
Cronbach's alpha	.74	.70	.75	.78	.82
Mean r_{pb}*	.14	.13	.15	.17	.19
SEM[i]	3.98	3.90	4.02	3.89	3.90
SEM at the pass mark	4.46	4.35	4.46	4.35	4.37
Decision consistency (uncorrected)[ii]	.91	.88	.91	.89	.91
Perceived fairness[iii]	65%	57%	59%	57%	53%
Pass mark	64.712	65.124	64.974	63.379	60.745
Effective pass mark	65	66	65	64	61
Pass rate	90.7%	89.6%	89.3%	86.0%	85.9%

[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 January 2020, the following exam development activities have taken place.

Validation

To renew the validation of items expiring from usability, a validation session was held with the EVC (see Table 32) remotely on May 5, 6, 11 and 12, 2020. Though originally planned to be an in person session, it was converted to remote because of COVID-19. During these 4 sessions, CKE 2 items were also validated, and that portion of the validation activity is reported in the technical report for that examination.

Note that scheduling precluded all identified members from being available for all 4 days; those marked as having participated attended the validation activity on at least 1 day.

Table 32: CHRL Examination Validation Committee – Validation

Member	Credential	Years of Relevant Experience	Start on EVC	Industry
✓ **Valentin Bachner**	CHRL	20–29	2019	Government & public agencies
✓ **Nadine Bellhouse**	CHRL	15–19	2019	Printing
✓ **Jennifer Borges**	CHRL	10–14	2017	Manufacturing
✓ **Annette Dhanasar**	CHRL	15–19	2017	Technology
✓ **Debbie Hynes**	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
✓ **Karen Pantaleo**	CHRL	20–29	2019	Healthcare / Consulting
✓ **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 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. For participants not able to join on the first day, they received individual training on the first day of their involvement. Each day, committee members were provided with approximately 50–70 items via a secure file share site, and then worked individually reviewing items through the day, submitting their appraisal and any suggested revisions to Wickett through the day. They were directed to make sure the items reflected current practice and were suitable to make decisions about who should receive the 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 184 items and case texts as suitable for the CHRL ELE, rejected 7 items, and shifted 0 items for eligibility in the CHRP ELE bank. Thirty-six 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 33.

Table 33: Employment Law Examination Blueprint Structural Variables

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

Content Weighting

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

Categories are:

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

Within each Category, the Topics are:

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

The full blueprinted list of Categories, Topics and Subtopics, along with their weighting, appears in Table 34.

Table 34: Employment Law Examination Blueprint Content Weights

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

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

Note: Reasonable ranges around the Topic weights are employed.

Cognitive Level

The cognitive level weights are based on Bloom's taxonomy. The purpose of this weighting is generally to ensure that an examination does not unintentionally over-focus on specific types of items, and to provide candidates with a range of items (in approximate proportion) that reflects the cognitive operations they must apply on the job. The weights appear in Table 35.

Table 35: 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.