

# Undercover Learners in Test Preparation Courses: A Test Security Perspective

Zhi Li, Claire Ewatski, Jayanti Banerjee

Paragon Testing Enterprises Inc. (research@paragontesting.ca)



## INTRODUCTION

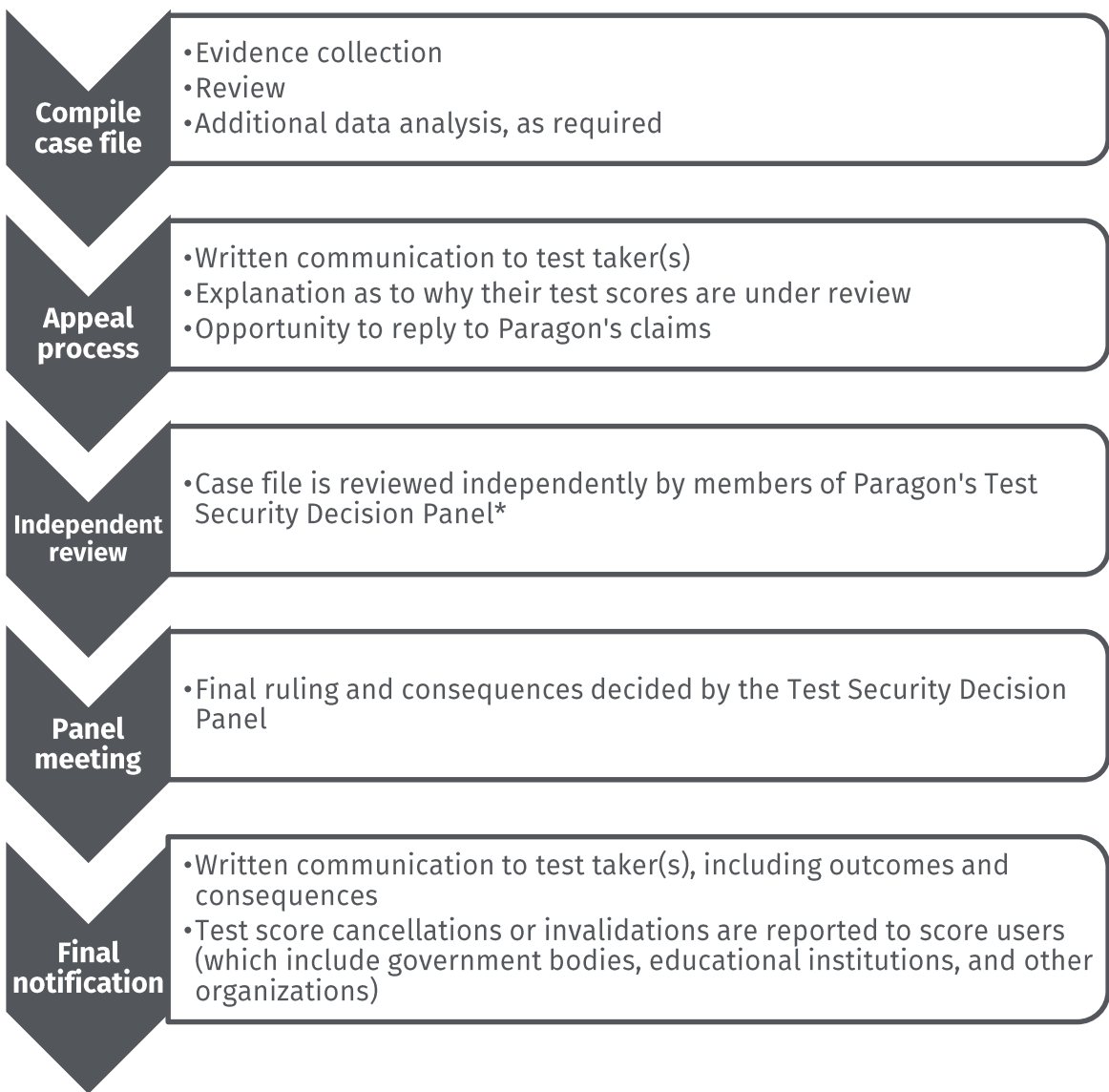
Research on test washback shows that test takers typically prepare for high-stakes tests by attending test preparation courses (Yu et al., 2017). These courses familiarize test takers with the test structure and reduce their anxiety. However, test preparation courses may also artificially inflate test scores and thus threaten the validity of the interpretation of the test result (Haladyna & Downing, 2005). Furthermore, unethical test preparation programs pose a threat to test security (Foster, 2015). This study reports on an investigation of two test preparation courses for a high-stakes English proficiency test by two undercover learners.

### Washback

- Washback can be defined as activities that teachers and learners engage in because of the test (typically to ensure that learners perform better on the test).
- This can be conceptualized as entirely value free: Students want to do well on a test and teachers would like to help their students do well on a test.
- Therefore, test takers and teachers will identify specific things that they can do which will boost test performance.
- This can include specific drills or practice tasks (cf. Alderson & Hamp Lyons, 1996).
- The extent to which the design of a test impacts test preparation activities is influenced by teacher beliefs (Watanabe, 1996), teacher expertise (Cheng, 1997), student motivation (Watanabe, 2001), and a multitude of systemic factors (Wall, 2000).
- When test preparation steps over the line from teaching language and language skills that will boost test performance to teaching a student simply how to pass the test, washback can become a test security concern (cf. Mehrens and Kaminski, 1989).
- Test preparation activities become even more problematic when they encourage fraudulent behaviors such as cheating and test piracy or theft, especially in computer-based tests (Foster, 2015).
- Limited attention has been directed to the test preparation programs that heavily, if not solely, focus on teaching the test (Yu et al., 2017).

### Test Security Investigation Process at Paragon

- Paragon's Test Security Unit is responsible for identifying and investigating any potential flag that may compromise the validity of a test taker's score. A fair and just process is followed:



*\*Paragon's Test Security Decision Panel consists of three rotating members of senior Paragon employees. Panelists are asked to review the case file independently and come to a decision before attending a panel meeting. Panelists deliberate during panel meetings to determine an overall case ruling.*

## METHODOLOGY



### The CELPIP test

- The CELPIP (Canadian English Language Proficiency Index Program) test is a computer-based language proficiency test designed to assess test takers' English abilities in a variety of everyday situations.
- Two versions of the CELPIP test are available:
  - CELPIP-General Test (consisting of reading, listening, speaking, and writing components) is used in applications for Canadian Permanent Residency and is also required by professional organisations such as provincial real estate councils.
  - CELPIP-General LS Test (consisting of listening and speaking components only) is used in applications for Canadian citizenship.

### Data collection and analysis

- A case study approach including the following data sources:
  - Case reports & email exchanges between the participants and the Test Security Unit
  - Semi-structured interview with the participants
- The interview data were transcribed verbatim.
- Open coding of the interview transcripts was conducted by two coders in the first round.
- The resultant codes were compared and consolidated, and then grouped for emerging themes.

### Undercover learners & test preparation programs

Participant A	• Speak English as their second language.	A face-to-face language test preparation course
Participant B	• Are proficient in the L1 of the learners in the test preparation courses.	An online test preparation course delivered through pre-recorded videos and online discussion forum

## EMERGING THEMES

### Teaching Approach

- Familiarization with test structure
- Drills
  - "... we've listened to some of the listening clips and they have some practice sheet and we do some exercise type of things." [L001]
- Test taking strategies
  - "And the instructor will explain, for example, why a key is wrong, is a right choice and why some distractors you can maybe eliminate like even before you listen to them ..." [L001]

### Fraudulent Practices

- Templated answers
  - "... you don't have to worry about writing and speaking once you're taking my course, because you just need to memorize what I'm saying. You know, like I have all these prepared model answers. So... yeah, you just need to memorize and you will get at least a like level five" [L002, paraphrasing the course tutor]
- Breach of copyright
  - "... our IPP materials will have practice test and test questions. They just reprint these materials and to distribute their students so their students do not purchase the materials." [L001]
- Item harvesting
  - "... he's like advertising, 'I- We do have a lot of test items here. And like those like testing companies in North America, they always re-use the test items in the past, so you just... need to follow me, like I have the items and the answers ready for you so you just have to trust me'." [L002]

## EMERGING THEMES – CONTINUED

### Test-Taker Awareness of Fraud

"Nobody felt like bad about this whole thing. Like- just- people just seem to know like- to know latest questions, and they- when somebody shares it they like appreciate them. But nobody says like, "This is not the right thing". Nobody." [L002]

"I feel, like, test-takers or the people who are trying to learn at the school, they did not do anything wrong. They just want to pass the test, like everything else is ready, right?" [L001]

### Undercover Learner Perspectives

- Fear of being found out
  - "I am not really good at saying like planned lies ..." [L001]
- Internal conflict over role
  - "I kinda felt "Am I like a... betrayer to the [L2] <laughs> community in Canada?". Yeah, I felt bad a little bit, but then again this is like so not right and I think I felt more rewarding doing this whole job." [L002]
- Empathy for test takers
  - "I think it will be nice if we actually, to some extent, sometimes look at it from test-takers' perspective and feel how they feel, to understand, like, their frustrations. ... They only want fair scores. In that sense, it is good for me to get a chance to sit there and see how they are doing and what do they feel when they take the test." [L001]

### Language as a Criteria for Immigration

"They do not really need to use English. So, they do not feel the need, for example, to spend one or two years or even longer to learn English." [L001]

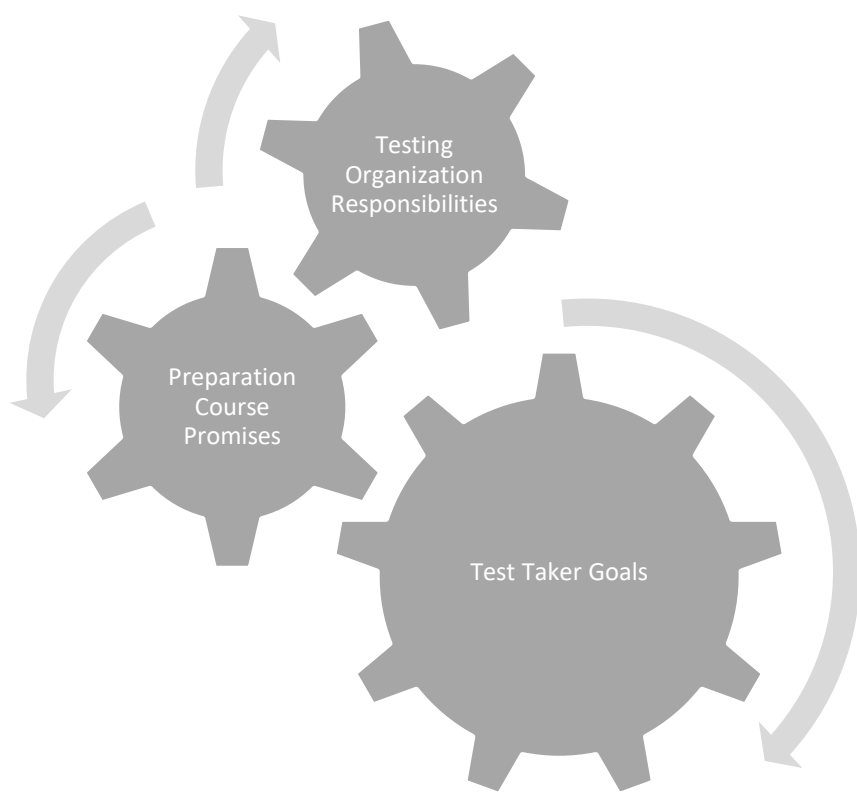
## REFLECTIONS

### What causes fraudulent behavior?

- Testing culture
- High-stakes nature of the test
- Policy decisions by organizations and governments

### Test fraud detection

- Is it fruitful to engage undercover learners?
- What are the ethical implications of engaging undercover learners?
- What is a testing organization's responsibility in helping test takers to understand the line between test preparation and test fraud?



## REFERENCES

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280–297.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38–54.
- Haladyna, T. M., & Downing, S. M. (2005). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17–27.
- Foster, D. (2015). Testing technology and its effects on test security. In F. Dragow (Ed.), *Technology and Testing: Improving Educational and Psychological Measurement* (pp. 235–254). Routledge.
- Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? *Educational Measurement: Issues and Practice*, 8(1), 14–22.
- Yu, G. et al. (2017). *Preparing for the speaking tasks of the TOEFL iBT® Test: An investigation of the journeys of Chinese test takers*. ETS Research Report Series.
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28(4), 499–509.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318–333.
- Watanabe, Y. (2001). Does the university entrance examination motivate learners? A case study of learner interviews. In *Trans-equator exchanges: A collection of academic papers in honour of Professor David Ingram* (pp. 100–110). Akita Association of English Studies.