

Chapter 3

Mitigating Risk: The Impact of a Diagnostic Assessment Procedure on the First-Year Experience in Engineering

Janna Fox, John Haggerty, and Natasha Artemeva

Abstract The global movement of students, the linguistic and cultural diversity of university classrooms, and mounting concerns about retention and program completion have prompted the increased use of post-entry diagnostic assessment, which identifies students at risk and provides them with early academic support. In this chapter we report on a multistage-evaluation mixed methods study, now in its sixth year, which is evaluating the impact of a diagnostic assessment procedure on the first-year experience, student engagement, achievement, and retention in an undergraduate engineering program. The diagnostic assessment procedure and concomitant student support are analyzed through the lens of Activity Theory, which views socio-cultural object-oriented human activity as mediated through the use of *tools*, both symbolic (e.g., language) and material (e.g., computers, pens). Changes in Activity Systems and their interrelationships are of central interest. In this chapter we report on changes resulting from modifications to the diagnostic assessment procedure that have increased its impact on the first-year experience by: (1) applying a disciplinary (rather than generic) assessment approach which was *fine grained* enough to trigger actionable academic support; (2) embedding the diagnostic assessment procedure within a required first-year engineering course, which increased the numbers of students who voluntarily sought support; and (3) paying increased attention to the development of social connections, which play an important role in student retention and success.

Keywords Activity theory • At-risk students • Diagnostic assessment • Engineering education • First-year experience • Higher education • Retention • Social connection

J. Fox (✉) • N. Artemeva

School of Linguistics and Language Studies, Carleton University, Ottawa, Canada
e-mail: Janna.Fox@carleton.ca; Natasha.Artemeva@carleton.ca

J. Haggerty

Department of Language and Literacy Education, University of British Columbia,
Vancouver, Canada
e-mail: john.haggerty@alumni.ubc.ca

1 Introduction

Pre-admission language proficiency testing has become a ubiquitous requirement for second language applicants to English-medium universities. Over the years, test users have tended to mistakenly interpret high scores on language proficiency tests as evidence of academic readiness (see, Fox et al. 2016), but if language proficiency alone were sufficient for successful academic engagement in an English-medium university, all English-speaking students would be successful. Clearly this is not the case.

In Canada, alongside the trend to ever larger numbers of international students, decades of immigration have contributed to increasing cultural and linguistic diversity in university classrooms (Anderson 2015). It is estimated that up to half of the students in Canada's schools speak a language other than English or French¹ as their first language (Fox 2015). In response to increasing cultural and linguistic diversity and greater concern about retention, universities have been developing a wide array of support services to address student needs. Typically, these services are generic and centralized, often with special attention directed at first-year undergraduates. For example, at the university where this study took place there are a number of such services available to students, including a writing tutorial service and a math tutorial service, amongst others.

Statistics on success in university suggest that first-year undergraduate students are the most likely to drop out or *stop out* (i.e., leave university for a period of time but return later) during the first term/s of their university programs (Browne and Doyle 2010). For this reason, *the first-year experience* has become a focus of much of the literature on university retention (e.g., Browne and Doyle 2010; Tinto 1993). Although there is evidence that generic support programs are helpful to students (Cheng and Fox 2008; Fox et al. 2014), at the university which is the site of the present study, statistics on retention rates suggest that such services have had relatively little impact on retention in the engineering program. For example, from 2007 through 2012, this program lost on average 16% of its students by the end of the first year; another 8% after second year (a 24% cumulative loss); and an additional 7% after third year (a 31% cumulative loss) (Office of Institutional Research and Planning 2014). Further, students within the engineering program were often taking the same course several times, and as a result, a number of them were not completing their programs in a timely manner. Many of the courses with the highest failure rates and repeat registrations were first- or second-year courses, which were part of the foundation or core engineering curriculum that all engineering students are required to complete successfully before specializing in an engineering discipline (e.g., civil, mechanical, electrical).

Within this context, and in light of mounting concerns in the Faculty of Engineering about retention, patterns of course enrollments (i.e., failure and

¹English and French are official languages of Canada and serve as mediums of instruction in Canadian universities.

repeated registration(s) in the same course), and delays in program completion, a post-entry diagnostic assessment procedure was implemented in 2010 for entering undergraduate students. The intent of the assessment was to identify at-risk students early in their program and provide them with individualized pedagogical support to mitigate risk of failure. Engineering professors, Teaching Assistants (TAs) and other stakeholders reported that, in addition to issues with language and writing, a number of first-year students lacked threshold concepts in mathematics, misunderstood instructions for assignments, underestimated how much work they needed to do on an on-going basis so as not to fall behind, and so on. It was evident that more than English language proficiency was creating risk for entering undergraduates in engineering.

In this chapter we report on a mixed methods research study examining the impact of the diagnostic assessment procedure, now in its sixth year, and guided by the following research question: Does a diagnostic assessment procedure, combining assessment with individual pedagogical support, improve the experience, retention, and achievement of entering undergraduate engineering students? This chapter reports on the impact of the diagnostic procedure on the first-year experience.

Before discussing the study itself, in the section which follows we discuss the diagnostic assessment procedure that has been developed for this university context within the theoretical framework and empirical research that informed it. We also provide background on the challenges and concerns that have characterized its implementation.

2 Evolution of the Diagnostic Assessment Procedure

Huhta (2008) distinguishes formative assessment from diagnostic assessment, submitting that they are on two ends of an assessment continuum. He suggests that formative assessment is rooted in on-going curricular and classroom concerns for feedback, teaching, and learning in *practice*, whereas diagnostic assessment is informed by theory and theoretical models of language and learning. However, in our view it is their connection, rather than what distinguishes them, that is of central concern. We use the phrase *diagnostic assessment procedure* to underscore that diagnostic assessment and concomitant pedagogical intervention are inseparable. We argue that a diagnostic assessment procedure cannot be truly diagnostic unless it is linked to feedback, intervention, and support. Alderson (2007) suggests this in his own considerations of diagnostic assessment:

... central to diagnosis must be the provision of usable feedback either to the learners themselves or to the diagnoser—the teacher, the curriculum designer, the textbook writer, and others. ... [T]he nature of [the] feedback, the extent to which it can directly or indirectly lead to improvements in performance or in eradicating the weaknesses identified, must be central to diagnostic test design. (p. 30)

In this study, the diagnostic assessment procedure has as its goal to: (1) identify entering students at-risk in the first year of their undergraduate engineering program; and (2) generate a useful learning profile (i.e., what Cai 2015 notes must lead to *actionable feedback*), which is linked to individually tailored (Fox 2009) and readily available academic support for the learning of a first-year engineering student.

The study is informed by sociocultural theory (Vygotsky 1987), which views knowledge as contextualized and learning as social (e.g., Artemeva and Fox 2014; Brown et al. 1989; Lave and Wenger 1991). From this perspective, both the assessment and the pedagogical interactions that it triggers are *situated* within and informed by the context and the community in which they occur (Lave and Wenger 1991). However, context is a complex and multi-layered construct, extending from micro to increasingly (infinitely) macro levels. Thus, early in the study we encountered what is known in the literature as the “frame problem” (Gee 2011a, p. 67, 2011b, p. 31), namely, to determine the degree of situatedness that would be most useful for the purposes of the diagnostic assessment procedure.

In the initial implementation of the diagnostic assessment, we operationalized what Read (2012) refers to as a general academic language proficiency construct, of relevance to university-level academic work. For our pilot in 2010, we drew three tasks from the Diagnostic English Language Needs Assessment (DELNA) test battery (see, Read 2008, 2012; or, the DELNA website: <http://www.delna.auckland.ac.nz/>). DELNA is a Post-Entry Language Assessment (PELA) procedure. As discussed in Read (2012), such PELA procedures operationalize the construct of academic English proficiency and draw on generic assessment materials of relevance across university faculties and programmes. Alderson (2007) argues that “diagnosis need not concern itself with authenticity and target situations, but rather needs to concentrate on identifying and isolating components of performance” (p. 29). Isolating key components of performance as the result of diagnosis provides essential information for structuring follow-up pedagogical support.

Of the three DELNA tasks that were used for the initial diagnostic assessment procedure, two were administered and marked by computer and tested academic vocabulary knowledge and reading (using multiple-choice and cloze-elide test formats). The third task tested academic writing with a graph interpretation task that asked test takers to write about data presented in a histogram. The task was marked by raters trained to use the DELNA rubric for academic writing. We drew two groups of raters, with language/writing studies or engineering backgrounds, and all were trained and certified through the DELNA on-line rater training system (Elder and von Randow 2008). However, while the language/writing raters tended to focus on language-related descriptors, the engineering raters tended to focus on content. Although inter-rater reliability was high (.94), there was a disjuncture between what the raters attended to in the DELNA rating scale and what they valued in the marking (Fox and Artemeva 2011). Further, some of the descriptors in the DELNA generic grid were not applicable to the engineering context (e.g., valuing length as opposed to concise expression; focusing on details to support an argument, rather than interpreting trends). The descriptors needed to map onto specific, actionable

pedagogical support summarized in the learning profile. Having descriptors which were not of relevance to engineering was unhelpful.

Following the research undertaken by Fox and Artemeva (2011), we recognized that considerations of disciplinarity were critical if interventions were to provide what Prior (1994) refers to as an “opportunity space for socialization into [the] discursive practices” (p. 489) of the discipline. As a result, we narrowed the *frame* to the context of first-year undergraduate engineering and began to operationalize an engineering-specific, *academic literacies construct* (Read 2015). The diagnostic assessment procedure aimed to create a safe support space which would allow new students (novices in the discipline) to begin to “display disciplinarity” (Prior 1994, p. 489). As Artemeva (2006) points out, novices typically go through a fairly slow process of acculturation before they can communicate what they know in ways acceptable to a new discipline. A Centre, staffed by peer mentors who were drawn from the pool of raters, was set up to provide support which would help facilitate a new student’s acculturation. This support space was much *safer* than the classroom, which is public, contested, and subject to grades or marks; safer, because the Centre was staffed by knowledgeable student mentors who were nonetheless external to courses and posed no threat.

3 Theoretical Framework

Our analysis of the diagnostic assessment procedure and concomitant student support discussed in this chapter are enriched through the use of Activity Theory (AT) (Engeström 1987; Engeström et al. 1999; Leont’ev 1981). Vygotsky (1987) argued that object-oriented human activity is always mediated through the use of signs and symbols (e.g., language, texts). One of Vygotsky’s students, Leont’ev (1981), later posited that *collective* human activity is mediated not only by symbolic, but also by material tools (e.g., pens, paper). Leont’ev (1981) represented human activity as a triadic structure, often depicted as a triangle, which includes a subject (i.e., human actors), an object (i.e., something or someone that is acted upon), and symbolic or material tools which mediate the activity. The subject has a motive for acting upon the object in order to reach an outcome. AT was further developed by Engeström (1987), who observed that “object-oriented, collective and culturally mediated human activity” (Engeström and Miettinen 1999, p. 9) is best modelled as an *activity system* (e.g., see Figs. 3.1 and 3.2). In other words, “an activity system comes into existence when there is a certain need . . . that can be satisfied by a certain activity” (Artemeva 2006, p. 44). As Engeström noted, multiple activity systems interact over time and space. In the present chapter, we consider the diagnostic assessment procedure as comprising two interrelated activity systems (Fig. 3.1): a diagnostic assessment activity and a pedagogical support activity.

The diagnostic assessment procedure prompts or instigates the development of the activity system of the undergraduate engineering student (Fig. 3.2), who voluntarily seeks support from peer mentors within the Centre in order to improve his or her academic achievement (i.e., grades, performance).

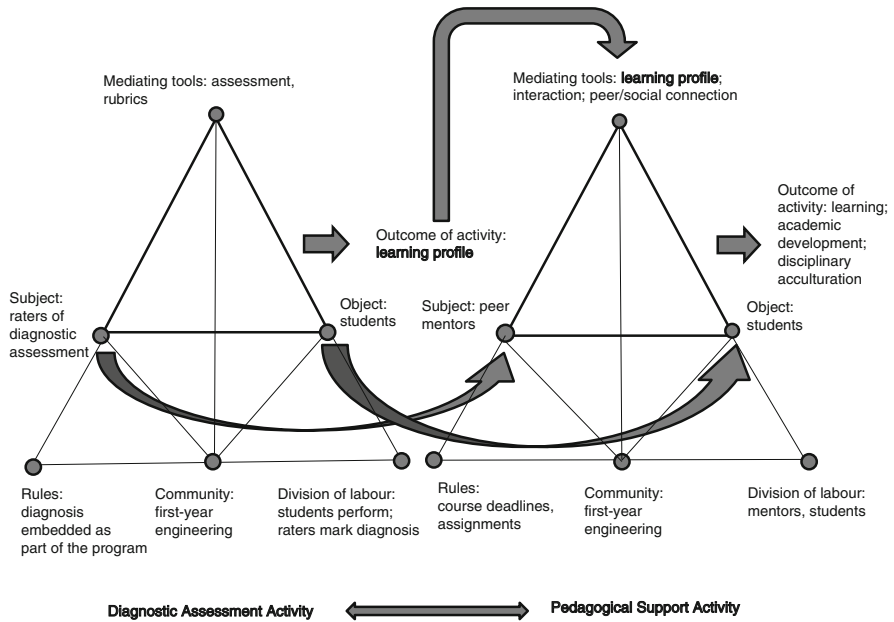


Fig. 3.1 Activity systems of the diagnostic assessment procedure in undergraduate engineering

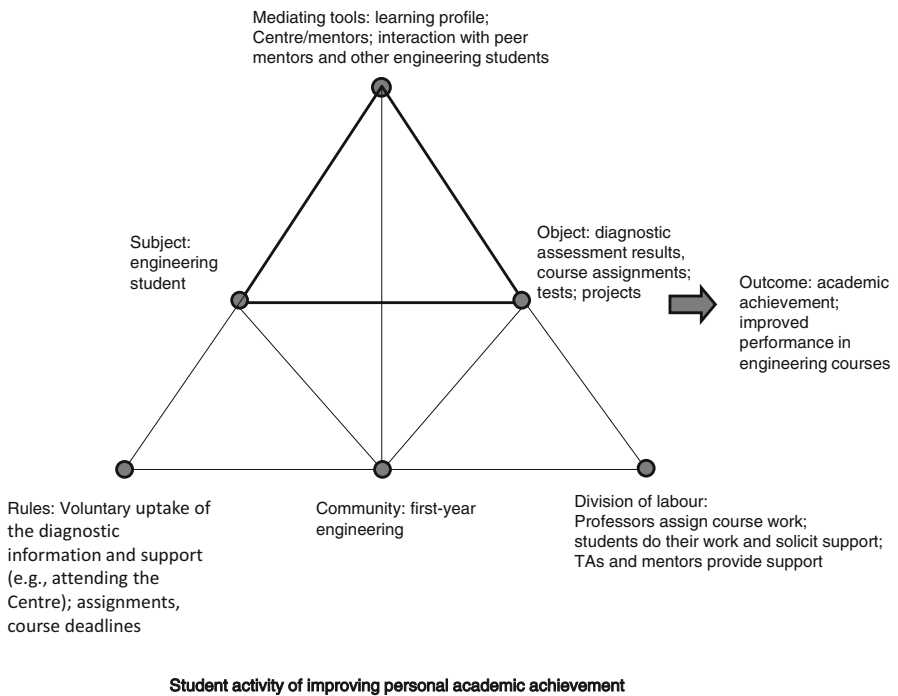


Fig. 3.2 The activity system of an undergraduate engineering student

It is important to note that the actors/subjects in the activity systems of the diagnostic assessment procedure (Fig. 3.1) are not the students themselves. As indicated above, the raters in the diagnostic assessment activity system become peer mentors in the pedagogical support activity. However, the students who use the information provided to them in the learning profile and seek additional feedback or pedagogical support bring into play another activity system (Fig. 3.2), triggered by the activities depicted in Fig. 3.1.

In our study, the activities presented in Figs. 3.1 and 3.2 are all situated within the community of undergraduate engineering, and, as is the case in any activity system (Engeström 1987), are inevitably characterized by developing tensions and contradictions. For example, there may be a contradiction between the mentors' and the students' activity systems because of contradictions in the motives for these activities. Mentors working with students within the Centre are typically motivated to support students' *long-term learning* in undergraduate engineering, whereas most students tend to have *shorter-term* motives in mind, such as getting a good grade on an assignment, clarifying instructions, or unpacking what went wrong on a test. According to AT, such tensions or contradictions between the mentors' activity system and the students' activity system are the site for potential change and development (Engeström 1987). Over time, a student's needs evolve and the student's motive for activity may gradually approach the motive of the mentors' activity. One of the primary goals of this study is to find evidence that the activity systems in Figs. 3.1 and 3.2 are aligning to ever greater degrees as motives become increasingly interrelated. Evidence that the activity systems of mentors and students are aligning will be drawn from students' increasing capability, and awareness of what works, why, and how best to communicate knowledge and understanding to others within their community of undergraduate engineering as an outcome of their use of the Centre. Increased capability and awareness enhance the first-year experience (Artemeva and Fox 2010; Scanlon et al. 2007), and ultimately influence academic success.

The diagnostic assessment procedure is also informed by empirical research which is consistent with the AT perspective described above. This research has investigated undergraduates' learning in engineering (e.g., Artemeva 2008, 2011; Artemeva and Fox 2010), academic and social engagement (Fox et al. 2014), and the process of academic acculturation (Cheng and Fox 2008), as well as findings produced in successive stages of the study itself (e.g., Fox et al. 2016). We are analyzing longitudinal data on an on-going basis regarding the validity of the at-risk designation and the impact of pedagogical support (e.g., retention, academic success/failure, the voluntary use of pedagogical support), and using a case study approach to develop our understanding of the phenomenon of *risk* in first-year engineering.

Having provided a brief discussion of the theoretical framework and on-going empirical research that inform the diagnostic assessment procedure, in the section below we describe the broader university context within which the procedure is situated and the issues related to its development and implementation.

4 Implementation of the Diagnostic Assessment: A Complex Balancing Act

The development and implementation of the diagnostic assessment procedure described here may be best described as a complex and challenging balancing act (Fox and Haggerty 2014). In maintaining this balance, there have been issues and tensions (see Fig. 3.3) regarding:

1. Assessment quality (e.g., tasks, scoring rubrics, rater consistency);
2. Pedagogical support; and
3. Presentation and marketing of the diagnostic assessment to key stakeholders (e.g., students, administrators, faculty, TAs).

In the section below, we discuss each of these dimensions of concern in relation to some of the questions that we have attempted to address over the 6 years of development, administration, and implementation of the diagnostic assessment procedure. This is not a comprehensive list by any means; rather, it provides an overview of the key questions that we needed to answer with empirical evidence garnered from on-going research.

Responses to the following questions have guided decision-making with regard to the three dimensions of concern:

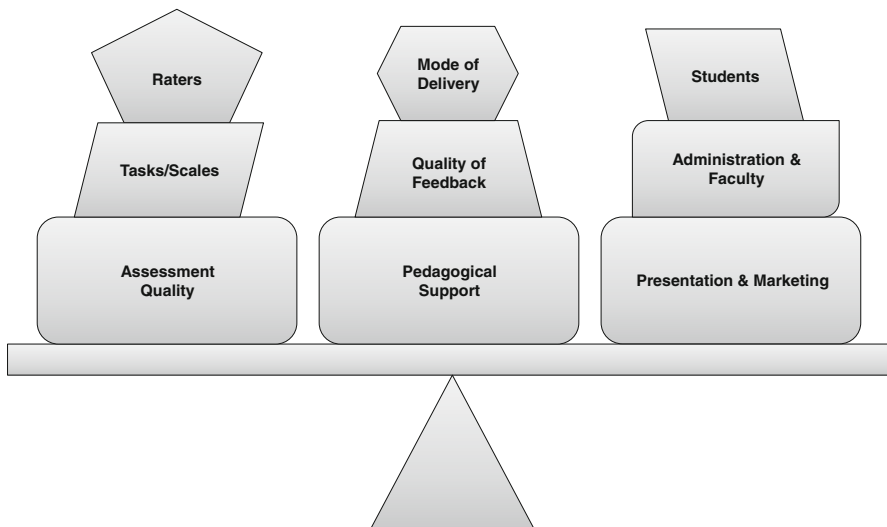


Fig. 3.3 Issues and tensions in diagnostic assessment: an ongoing balancing act

4.1 *Assessment Quality*

When a mandate for an assessment procedure had been established, and it had been determined that a test or testing procedures would best address the mandate, the critical first question is ‘do we know what we are measuring?’ (Alderson 2007, p. 21). As Alderson points out, “Above all, we need to clarify what we mean by diagnosis ... and what we need to know in order to be able to develop useful diagnostic procedures” (p. 21). Assessment quality depends on having theoretically informed and evidence-driven responses to each of the following key questions:

- 4.1.1 Is the construct well-enough understood and defined to warrant its operational definition in the test?
- 4.1.2 Are the items and tasks meaningfully operationalizing the construct? Do they provide information that we can use to shape pedagogical support?
- 4.1.3 Is the rating scale sufficiently detailed to provide useful diagnostic information?
- 4.1.4 Are the raters consistent in their interpretation of the scale?
- 4.1.5 Does the rating scale and the resulting score or scores provide sufficient information to trigger specific pedagogical interventions?

4.2 *Pedagogical Support*

Once the results of a diagnostic assessment are available, there are considerable challenges in identifying the most effective type of pedagogical support to provide. The other contributors to this volume have identified the many different approaches taken to provide academic support for students, especially for those who are identified as being at-risk. For example, in some contexts academic counsellors are assigned to meet with individual students and provide advice specific to their needs (e.g. Read 2008). In other contexts, a diagnosis of *at-risk* triggers a required course or series of workshops.

Determining which pedagogical responses will meet the needs of the greatest number of students and have the most impact on their academic success necessitates a considerable amount of empirical research, trial, and (unfortunately) error. Tensions occur, however, because there are always tradeoffs between what is optimal support and what is practical and possible in a given context (see Sect. 4.3 below).

Research over time helps to identify the best means of support. Only with time is sufficient evidence accumulated to support pedagogical decisions and provide evidence-driven responses to the following questions.

- 4.2.1 Should follow-up pedagogical support for students be mandatory? If so, what should students be required to do?
- 4.2.2 Should such support be voluntary? If so, how can the greatest numbers of students be encouraged to seek support? How can students at-risk be reached?

- 4.2.3 What type of support should be provided, for which students, when, how, and by whom?
- 4.2.4 How precisely should information about test performance and pedagogical support be communicated to the test taker/student? What information should be included?
- 4.2.5 How do we assess the effectiveness of the pedagogical support? When should it begin? How often? For what duration?
- 4.2.6 What on-going research is required to evaluate the impact of pedagogical interventions? Is there evidence of a change in students' engagement and their first-year experience?
- 4.2.7 What evidence do we need to collect in order to demonstrate that interventions are working to increase overall academic retention and success?

4.3 Presentation and Marketing

Diagnostic assessment and subsequent pedagogical responses are time-consuming and costly. One of the key tensions in considerations of diagnostic assessment procedures is the need to provide evidence that additional cost is warranted by the positive benefits that result.

Presentation is critical to persuading university administrators who control budgets that offering a diagnostic assessment to students early in their undergraduate programs will promote retention and academic success. From the beginning, it is important to collect evidence on an on-going basis to demonstrate how diagnosis and intervention are making a difference.

In some contexts, students at-risk are required to undertake a course, participate in workshops, or receive counselling as a result of the diagnosis. Although some universities may provide the necessary funding to cover these additional costs, in many instances students who are required to take an extra course must pay for it themselves or pay a small fee (in the case of additional, mandatory workshops). In the context of the diagnostic assessment procedure which is the focus of the present chapter, students were not required to use the diagnostic information. Students' use or follow-up was entirely voluntary. Thus, presentation and marketing of the usefulness of the diagnostic assessment procedure to students was a key concern and led to many challenges which were addressed through evidence-driven responses to the following key questions:

- 4.3.1 How can we best encourage students to follow-up on the diagnostic information provided by the assessment?
- 4.3.2 What evidence should be collected in order to persuade administrators that the assessment procedure is having an impact on retention and academic success?
- 4.3.3 How, when, and where should pedagogical support be delivered?
- 4.3.4 Who should provide pedagogical support? What funding is available for this cost?

4.3.5 How should we recruit and train personnel to provide effective pedagogical support?

4.3.6 Who should monitor their effectiveness?

4.3.7 How clear is the tradeoff between cost and benefit?

4.3.8 How can we tailor the diagnostic assessment procedure and pedagogical follow-up to maximize practicality, cost-effectiveness, and impact?

Given the complex balancing act of developing and implementing a diagnostic assessment procedure, we have limited our discussion in the remainder of this chapter to two changes that have occurred since the assessment was first introduced in 2010. These evidence-driven changes were perhaps the most significant in improving the quality of the assessment itself and increasing the overall impact of the assessment procedure. The two changes that were implemented in 2014 are:

- Embedding diagnostic assessment within a mandatory, introductory engineering course; and,
- Providing on-going pedagogical support during the full academic year through the establishment of a permanent support Centre for engineering students.

These changes were introduced as a result of research studies, occurring at different phases and multiple stages of development and implementation of the procedure, informed by the theoretical framework, and guided by the questions listed above and the overall research question: Does a diagnostic assessment procedure, combining assessment with individual pedagogical support, improve the first-year experience, achievement, and retention of undergraduate engineering students? If so, in what ways, how, and for whom? In this chapter we focus on evidence which relates the two changes in 2014 to differences in the first-year experience.

5 The Current Study: Two Significant Changes

5.1 Method

The two key changes to the diagnostic assessment procedure, namely, embedding it in a required course and providing a permanent Centre for support, were supported by initial results of the longitudinal mixed-methods study which is exploring the effectiveness of this diagnostic assessment procedure by means of a *multistage-evaluation design* (Creswell 2015). As Creswell notes, “[t]he intent of the multistage evaluation design is to conduct a study over time that evaluates the success of a program or activities implemented into a setting” (p. 47). It is multistage in that each phase of research may, in effect, constitute many stages (or studies). These stages may be qualitative (QUAL or qual, depending upon dominance), quantitative (QUAN or quan, depending upon dominance) or mixed methods, but taken together they share a common purpose. Figure 3.4 provides an overview of the research design and includes information on the studies undertaken within phases of the

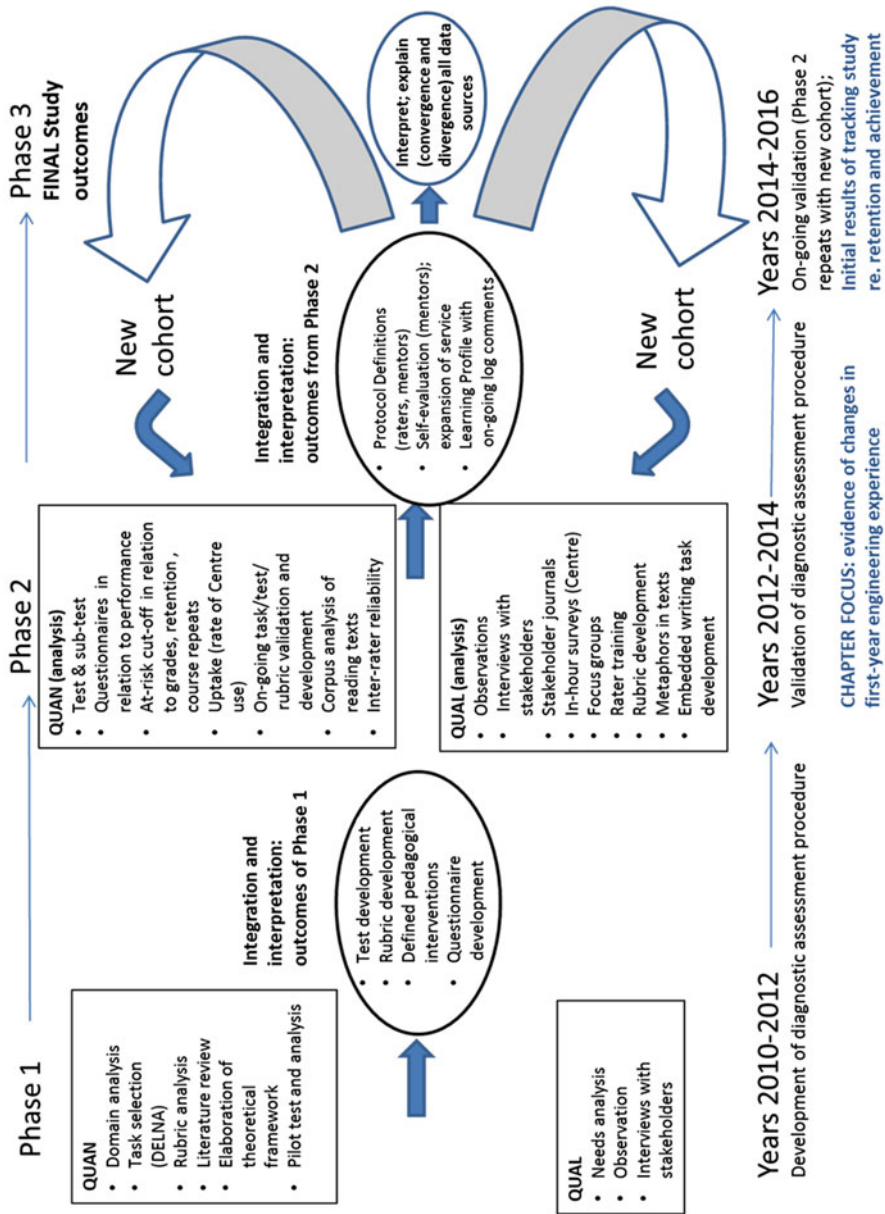


Fig. 3.4 Mixed methods design of the diagnostic assessment procedure

research; how the qualitative and quantitative strands were integrated; and the relationship of one phase of the research to the next from 2010 to 2016.

Engaging in what is essentially a ‘program-of-research’ approach is both complex and time-consuming. Findings are reported concurrently as part of the overall research design. For example, within the umbrella of the shared purpose for this longitudinal study, Fox et al. (2016) report on a Phase 1 research project investigating the development of a writing rubric, which combines generic language assessment with the specific requirements of writing for undergraduate engineering purposes. As discussed above, the new rubric was the result of tensions in the diagnostic assessment procedure activities (Fig. 3.1) in Phase 1 of the study. The generic DELNA grid and subsequent learning profile contradicted some of the expectations of engineering writing (e.g., awarded points for length, anticipated an introduction and conclusion, did not value point form lists). Increasing the engineering relevance of the rubric also increased the diagnostic feedback potential of the assessment procedure and triggered differential pedagogical interventions. In other words, tensions in the original activity system led to the development of a more advanced activity system (Engeström 1987), which better identifies and isolates components of performance (Alderson 2007). In another Phase 1 project, McLeod (2012) explored the usefulness of diagnostic feedback in a case study which tracked the academic acculturation (Artemeva 2008; Cheng and Fox 2008) of a first-year undergraduate engineering student over the first year of her engineering program. She documented the interaction of diagnostic assessment feedback, pedagogical support, and the student’s exceptional social skill in managing risk, as the student navigated through the challenges of first-year engineering. McLeod’s work further supports the contention that social connections within and across a new disciplinary community directly contribute to a student’s academic success.

Another component of the multistage evaluation design includes a qualitative case study in Phase 2 with academic success in engineering being the phenomenon of interest. The case study is drawing on semi-structured interviews with first-year and upper-year engineering students, TAs, professors, peer mentors, administrators, etc. to investigate the impact of the diagnostic assessment from different stakeholder perspectives. There is also a large-scale quantitative study in Phase 2 which is tracking the academic performance of students who were initially identified as being at-risk and compares their performance with their more academically skilled counterparts using various indicators of student academic success (e.g., course withdrawals, course completions and/or failures, drop-out rates, and use of the Centre which provides pedagogical support).

As indicated above, the design is recursive in that the analysis in Phase 2 will be repeated for a new cohort of entering engineering students and the results will further develop the protocols designed to support their learning. The current study will be completed in 2016. On-going in-house research is now the mandate of the permanent Centre (see Fig. 3.4, Phase 3).

Table 3.1 Summary of participants by stakeholder group (2010–2015)

Stakeholder group	Pilot test	External to course	External to course	Embedded in course	Embedded in course
Year	2010–2011	2011–2012	2012–2013	2013–2014	2014–2015
Engineering students	160	489	518	971	1,014 ^a
Professors/instructors	5	3	5	5	5
Peer mentors		11	5	7	11 ^b
Trained raters	15	8	4	6	6
Administrators (project coordinator; dean; associate dean, etc.)	1	1	3	3	3
Total	181	512	535	992	1,039

^a972 students were tested in September 2014; 42 were tested in January 2015

^b3 were upper-year students in writing/language studies; 8 engineering

5.2 Participants

In large-scale mixed-methods studies which utilize a multistage-evaluation design, the numbers and types of participants tend to fluctuate continuously over time (Table 3.1).

As Creswell (2015) points out and as the name implies, this advanced type of mixed methods design is comprised of many stages within multiple phases, all of which collectively support a sustained line of inquiry, (e.g., needs assessment, development of a conceptual framework, field testing of prototypes). Each phase in the inquiry may feature “rigorous quantitative and qualitative research methods” (p. 3) or mixed methods, but the core characteristic of such a study is the *integration* (Fig. 3.4) of findings in relation to the overall intent of the research. What distinguishes a multistage evaluation design is that “integration consists of expanding one stage into other stages over time” (p. 7).

Although each stage of research responds to specific questions which dictate a particular sampling strategy (Creswell 2015), and the large-scale tracking study is considering the whole population of undergraduates in engineering, in Table 3.1 we provide an overview of participants from 2010 (when the initial assessment consisting of three DELNA tasks was first pilot tested) to 2015.

5.3 Findings and Discussion

As indicated in the introduction to this chapter, from the beginning there were two overarching concerns to address in evaluating the impact of the diagnostic assessment procedure within this engineering context:

- Development of a post-entry diagnostic assessment which would effectively identify undergraduate engineering students at-risk early in their first term of study; and

- Provision of effective and timely academic support for at-risk students, as well as any other first-year engineering students, who wanted to take advantage of the support being offered.

Initially, although students were directly encouraged to take the diagnostic assessment, their participation and use of the information provided by the assessment as well as follow-up feedback and post-entry support was *voluntary*. There were no punitive outcomes (e.g., placement in a remedial course; required attendance in workshops; reduction in course loads and/or demotion to part-time status). Students received feedback and advice on their diagnostic assessment results by email a week after completing the assessment. Their performance was confidential (neither their course professors nor the TAs assigned to their courses were informed of their results). The emails urged students to drop in for additional feedback at a special Centre to meet with other, upper year students (in engineering and writing/language studies) and get additional feedback, information, and advice on how to succeed in their engineering courses.

In 2014, two critical changes occurred in the delivery of the diagnostic assessment procedure which dramatically increased its impact. These two changes, which are the focus of the findings below, were the cumulative result of all previous stages of research, and as indicated above, have had to date the largest impact on the quality of the diagnostic assessment procedure.

Each of the key changes implemented in 2014 is discussed separately in relation to the findings which informed the changes.

5.3.1 Evidence in Support of Embedding the Diagnostic Assessment Procedure in a Mandatory First-Year Engineering Course

In Phase 1 of the study (2011–2012), 489 students (50% of the first-year undergraduate engineering cohort) were assessed with three of DELNA's diagnostic tasks leased from the University of Auckland (DELNA's test developer). The DELNA tasks were administered during *orientation week* – the week which precedes the start of classes in a new academic year and introduces new students to the university. Students were informed of their results by email and invited on a voluntary basis to meet with peer mentors to receive pedagogical support during the first months of their engineering program. A Support Centre for engineering students was set up during the first 2 months of the 2011–2012 term. It was staffed by upper-year engineering and writing/language studies students who covered shifts in the Centre from Monday through Friday, and who had previously rated the DELNA writing sub-test.

During the 6 weeks the Centre was open, only 12 (2%) of the students sought feedback on their diagnostic assessment results. The students tended to be those who were outstanding academically and would likely avail themselves of every opportunity to improve their academic success, or students who wanted information on an assignment. Only 3 of the 27 students who were identified as at-risk (11%) visited the Centre for further feedback on their results and took advantage of the

pedagogical support made available. At the end of the academic year, ten of the at-risk group had dropped out or were failing; seven were borderline failures; and, ten were performing well (including the three at-risk students who had sought feedback).

In 2012–2013, 518 students (70 % of cohort) were assessed, but only 33 students (4 %) voluntarily followed-up on their results. However, there was evidence that three of these students remained in the engineering program because of early diagnosis and pedagogical intervention by mentors in the Centre. Learning of the success of these three students, the Dean of Engineering commented, “Retaining even two students pays for the expense of the entire academic assessment procedure.”

In 2013–2014, the DELNA rating scale was adapted to better reflect the engineering context. This *hybrid* writing rubric (see Fox et al. 2016) improved the *grain size* or specificity of relevant information provided by the rubric and enhanced the quality of pedagogical interventions. Further, DELNA’s graph interpretation task was adapted to represent the engineering context. Graphic interpretation is central to work as a student of engineering (and engineering as a profession as well). However, when the DELNA graph task was vetted with engineering faculty and students, they remarked that “engineers don’t do histograms”. This underscored the importance of disciplinarity (Prior 1994) in this diagnostic assessment procedure.

It became clear that many of the versions of the generic DELNA graph task were more suited for social science students than for engineering students, who most frequently interpret *trends* with line graphs. In order to refine the diagnostic feedback elicited by the assessment and shape pedagogical interventions to support students in engineering, it became essential that engineering content, tasks, and conventions be part of the assessment. Evidence suggested that the *frame* of general academic language proficiency (Read 2015) was too broad; decreasing the frame size and situating the diagnostic assessment procedure with engineering text, tasks, and expectations of performance also increased both the overall usefulness of the assessment (Bachman and Palmer 1996) as well as the relevance and specificity (grain size) of information included in the learning profiles of individual students. More specificity in learning profiles also increased the quality and appropriacy of interventions provided to students in support of their learning. From the perspective of Activity Theory, the mentors were increasingly able to address the students’ motives: to use the learning profiles as a starting point; to mediate activity in relation to the students’ motives to improve their performance in their engineering classes or achieve higher marks on a lab report or problem set. The increased relevance of the mentors’ feedback and support suggests increased alignment between the activity systems of mentors and students (Figs. 3.1 and 3.2).

Domain analysis which investigated undergraduate engineering supported the view that engineering students might be at-risk due to more than academic language proficiency. For example, some students had gaps in their mathematics background while others had difficulty reading scientific texts. Still others faced challenges in written expression required for engineering (e.g., concise lab reports; collaborative or team writing projects). Importantly, a number of entering students, who were

deemed at-risk when the construct was refined to reflect requirements for academic success in engineering, were first-language English speakers. Thus, as the theory and research had suggested, a disciplinary-specific approach would potentially have the greatest impact in supporting undergraduates in engineering.

As early as 2012–2013, we had pilot tested two line graphs to replace the generic DELNA graphs. The new graphs illustrated changes in velocity over time in an acceleration test of a new vehicle. However, the graphs proved to be too difficult to interpret, given that they appeared on a writing assessment without any supporting context. Convinced that it was important to provide a writing task that better represented writing in engineering, in 2013 we also piloted and then administered a writing task that was embedded in the first lecture of a mandatory course, which all entering engineering students are required to take regardless of their future discipline (e.g., mechanical, aerospace, electrical engineering). During the first lecture, the professor introduced the topic, explained its importance, showed a video that extended the students' understanding of the topic, and announced that in the following class, the students would be asked to write about the topic by explaining differences in graphs which illustrated projected versus actual performance. Students were invited to review the topic on YouTube and given additional links for readings, should they choose to access them.

In 2014–2015, using the same embedded approach, we again administered the engineering graph task to 1014 students (99% of the cohort). As in 2013, students wrote their responses to the diagnostic assessment in the second class of their required engineering course. The writing samples were far more credible and informative than had been the case with the generic task, which was unrelated to and unsupported by their academic work within the engineering program. The information produced by the hybrid rubric provided more useful information for peer mentors. Other diagnostic tasks were added to the assessment including a reading from the first-year chemistry textbook in engineering, with a set of multiple choice questions to assess reading comprehension, and a series of mathematics problems which represented foundational mathematics concepts required for first-year.

Thus, the initial generic approach evolved into a diagnostic procedure that was embedded within the university discipline of engineering and operationalized as an academic literacy construct as opposed to a language proficiency construct. The embedded approach is consistent with the theory that informed the study. As Figs. 3.1 and 3.2 illustrate, activities are situated within communities characterized by internal rules and a division of labour.

In the literature on post-entry diagnostic assessment, an embedded approach was first implemented at the University of Sydney in Australia as Measuring the Academic Skills of University Students (MASUS) (Bonanno and Jones 2007; Read 2015). Like the diagnostic procedure that is the focus of the present study, MASUS:

- operationalizes an academic literacy construct,
- draws on materials and tasks that are representative of a discipline,
- is integrated with the teaching of courses within the discipline, and
- is delivered within a specific academic program.

In December 2014, data from field notes collected by peer mentors indicated a dramatic increase in the number of students who were using the Centre (see details below). In large part, the establishment of a permanent support Centre, staffed with upper-year students in both engineering and in writing/language studies filled a gap in disciplinary support that had been evident for some time. The change in the mediational tools in the support activity system (i.e., to a permanent Centre), embedded within the context of the first-year required course in engineering, has also had an important impact on student retention and engagement.

5.3.2 Evidence Supporting a Permanent Place for Engineering Support: The Elsie MacGill Centre

In the context of voluntary uptake (Freadman 2002), where the decision to seek support is left entirely to the student, one of the greatest challenges was *reaching* students at-risk. As McLeod (2012) notes, such students are at times fearful, unsure, unaware, or unwilling to approach a Centre for help – particularly at the beginning of their undergraduate program. Only those students with exceptional social networking skills are likely to drop-in to a support Centre in the first weeks of a new year. These students manage a context adeptly (Artemeva 2008; Artemeva and Fox 2010) so that support works to their advantage (like the at-risk student who was the focus of McLeod’s research).

From 2010 to 2012, the Centre providing support was open during the first 2 months of the Fall term (September–October) and was located in a number of different sites (wherever space was available). When the Centre closed for the year, interviews with engineering professors and TAs, instructors in engineering communications courses, and upper-year students who had worked in the Centre suggested the need for pedagogical support was on-going. Of particular note was the comment of one TA in a first-year engineering course who recounted an experience with a student who was failing. She noted, “I had no place to send him. He had no place to go.” This sentiment was echoed by one of the engineering communications instructors who, looking back over the previous term, reported that one of her students had simply needed on-going support to meet the demands of the course. However, both she and her TA lamented their inability to devote more time to this student: “He was so bright. I could see him getting it. But there was always a line of other students outside my office door who also needed to meet with me. I just couldn’t give him enough extra time to make a difference.” Again, the type of support the student needed was exactly that which had been provided by the Centre. It was embedded in the context of engineering courses, providing on-going relevant feedback on engineering content, writing, and language.

As a result of evidence presented to administrators that the diagnostic assessment procedure and concomitant pedagogical support were having a positive impact, in 2013 a permanent space in the main engineering building was designated and named

the Elsie MacGill Centre² (by popular vote of students in the engineering program). It was staffed for the academic year by 11 peer mentors, 8 upper-year students from engineering and 3 from language/writing studies. In addition, funding was made available for on-going research to monitor the impact of the assessment procedure and pedagogical support.

From an Activity Theory perspective, the engagement of engineering students in naming, guiding, and increasingly using the Centre is important in understanding the evolution of the initial activities (see Figs. 3.1 and 3.2). As students increasingly draw on the interventions provided by mentors within the Centre (Fig. 3.2), motives of the two activities become more aligned and coherent; the potential for the novice's participation in the community of undergraduate engineering is increased because motives are less likely to conflict (or at least will be better understood by both mentors and students). As motives driving the activities of mentors and students increasingly align, the potential for positive impact on a student's experience, retention and academic success is also increased. Evidence of increasingly positive impact was gathered from a number of sources.

From September to December 2014, the peer mentors with an engineering background recorded approximately 135 mentoring sessions (often with one student, but also with pairs or small groups). However, the engineering peer mentors did not document whether students seeking support had been identified as at-risk.

During the same 3 month period, three students in the at-risk group made repeated visits (according to the log maintained by the writing/language studies peer mentors). However, not only at-risk students were checking in at the Centre and asking for help. There were 46 other students who used the pedagogical support provided by the writing/language studies peer mentors in the Elsie Centre (as it is now popularly called). In total, approximately 184 first-year students (19% of cohort) sought pedagogical support in the first 3 months of the 2014–2015 academic year, and the number has continued to grow. Peer mentors reported that there were so many students seeking advice that twice during the first semester they had to turn some students away.

Increasingly, second-year students were also seeking help from the Elsie Centre. The majority were English as a Second Language (ESL) students who were struggling with challenges posed by a required engineering communications course. It was agreed, following recommendations of the engineering communications course instructors, that peer mentors would work with these students as well as all first-year students in the required (core) engineering course. In January 2015, one of the engineering communications instructors, with the support of the Elsie Centre, began awarding 1% of a student's mark in the communications course for a visit and consultation at the Elsie Centre.

²Elsie MacGill was the first woman to receive an Electrical Engineering degree in Canada and the first woman aircraft designer in the world. She may be best known for her design of the Hawker Hurricane fighter airplanes during World War II. Many credit these small and flexible airplanes for the success of the Allies in the Battle of Britain. Students within the engineering program voted to name the Centre after Elsie MacGill.

Consistent with the theoretical framework informing this research, situating the diagnostic assessment procedure within a required engineering course has made a meaningful difference in students' voluntary uptake (Freadman 2002) of pedagogical support. As discussed above, in marketing the diagnostic assessment procedure to these engineering students, it was critical to work towards student ownership. Findings suggest that the students' increased ownership is leading to an important change in how students view and participate in the activity of the Centre. The engineering students in the 2014–2015 cohort seem to view the 'Elsie Centre' as an integral part of their activity system. As one student, who had just finished working on a lab report with a writing/language studies mentor, commented: "That was awesome. I'm getting to meet so many other students here, and my grades are getting better. When I just don't get it, or just can't do it, or I feel too stressed out by all the work...well this place and these people have really made a difference for me."

The Elsie Centre mentors have also begun to offer workshops for engineering students on topics and issues that are challenging, drawing on the personal accounts of the students with whom they have worked. The mentors have also undertaken a survey within the Centre to better understand what is working with which students and why. The survey grew out of the mentors' desire to elicit more student feedback and examine how mentors might improve the quality and impact of their pedagogical support. The activity system of the diagnostic assessment procedure (Fig. 3.1) is evolving over time, informed by systematic research, self-assessment, and the mentors' developing motive to be more effective in supporting more students. In other words, a more advanced activity system is emerging (Engeström 1987) which allows for further alignment in the activity systems of the mentors and the first-year students (Figs. 3.1 and 3.2).

6 Conclusion

Although the final figures for the 2014–2015 academic year are not yet available, there is every indication that the two changes made to the diagnostic assessment procedure, namely embedding the assessment in the content and context of a first-year engineering course, and setting up a permanent support Centre named and *owned* by these engineering students, have greatly increased its impact. Students are more likely to see the relevance and usefulness of diagnostic feedback and pedagogical support when it relates directly to their performance in a required engineering course. The Centre is open to *all* first-year students, and students of all abilities are using it. As a result, the Centre does not suffer from the stigma of a mandatory (e.g., remedial) approach. Increased engagement and participation by students is evidence of a growing interconnectedness that is shaping students' identities as members of the undergraduate engineering community. Lave and Wenger (1991) and Artemeva (2011) discuss the development of a *knowledgeably skilled identity* as an outcome of a novice's learning to engage with and act with confidence within a community. The development of this new academic identity (i.e., functioning

effectively as a student in engineering) contributes to the novice student's ability to act, increases the student's potential to learn (Artemeva 2011), and enhances their first-year experience. However, because use of the Centre remains voluntary, it remains an open question whether the Centre is reaching a sufficient number of students at-risk and is a focus of on-going research.

One of the most important outcomes of the two changes to the diagnostic assessment procedure was underestimated in its initial design. The Elsie Centre is facilitating the development of social connections within the engineering program, as students new to undergraduate engineering increase their sense of connection and community through interaction with peer mentors and their classmates in this newly created learning space. The findings from this study are consistent with the Vygotskian (1987) notions of knowledge and learning as situated and social. Empirical research on engagement identifies both academic and social considerations as key variables in predicting success in university study (e.g., Fox et al. 2014; Scanlon et al. 2007). Indeed, social connections that are fostered by interactions in the Elsie MacGill Centre may often be as important as academic connections in terms of enhanced first-year experience, retention, and levels of academic success. In the coming years, we will evaluate this relationship further with regard to the results of the tracking study, which will report on retention and academic success for new cohorts of entering undergraduate engineering students who have participated in the diagnostic assessment procedure described in this chapter.

References

- Alderson, J. C. (2007). The challenges of (diagnostic) testing: Do we know what we are measuring? In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 21–39). Ottawa: University of Ottawa Press.
- Anderson, T. (2015). Seeking internationalization: The state of Canadian higher education. *Canadian Journal of Higher Education*, 45(4), 166–187.
- Artemeva, N. (2006). Approaches to learning genres: A bibliographical essay. In N. Artemeva & A. Freedman (Eds.), *Rhetorical genre studies and beyond* (pp. 9–99). Winnipeg: Inkshed Publications.
- Artemeva, N. (2008). Toward a unified theory of genre learning. *Journal of Business and Technical Communication*, 22(2), 160–185.
- Artemeva, N. (2011). “An engrained part of my career”: The formation of a knowledge worker in the dual space of engineering knowledge and rhetorical process. In D. Starke-Meyerring, A. Paré, N. Artemeva, M. Horne, & L. Yousoubova (Eds.), *Writing in knowledge societies* (pp. 321–350). Fort Collins: The WAC Clearinghouse and Parlor Press. Available at <http://wac.colostate.edu/books/winks/>
- Artemeva, N., & Fox, J. (2010). Awareness vs. production: Probing students' antecedent genre knowledge. *Journal of Business and Technical Communication*, 24(4), 476–515.
- Artemeva, N., & Fox, J. (2014). The formation of a professional communicator: A sociorhetorical approach. In V. Bhatia & S. Bremner (Eds.), *The Routledge handbook of language and professional communication* (pp. 461–485). London: Routledge.
- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

- Bonanno, H., & Jones, J. (2007). *The MASUS procedure: Measuring the academic skills of University students. A resource document*. Sydney: Learning Centre, University of Sydney. Retrieved from http://sydney.edu.au/stuserv/documents/learning_centre/MASUS.pdf
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Browne, S., & Doyle, H. (2010). *Discovering the benefits of a first year experience program for under-represented students: A preliminary assessment of Lakehead University's Gateway Program*. Toronto: Higher Education Quality Council of Ontario.
- Cai, H. (2015). *Producing actionable feedback in EFL diagnostic assessment*. Paper delivered at the Language Testing Research Colloquium (LTRC), Toronto, 18 Mar 2015.
- Cheng, L., & Fox, J. (2008). Towards a better understanding of academic acculturation: Second language students in Canadian universities. *Canadian Modern Language Review*, 65(2), 307–333.
- Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Los Angeles: Sage.
- Elder, C., & von Randow, J. (2008). Exploring the utility of a web-based English language screening tool. *Language Assessment Quarterly*, 5(3), 173–194.
- Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research*. Helsinki: Orienta-Konsultit Oy.
- Engeström, Y., & Miettinen, R. (1999). Introduction. In Y. Engeström, R. Miettinen, & R. I. Punamäki (Eds.), *Perspectives on activity theory* (pp. 1–16). Cambridge: Cambridge University Press.
- Engeström, Y., Miettinen, R., & Punamäki, R. I. (Eds.). (1999). *Perspectives on activity theory*. Cambridge: Cambridge University Press.
- Fox, J. (2009). Moderating top-down policy impact and supporting EAP curricular renewal: Exploring the potential of diagnostic assessment. *Journal of English for Academic Purposes*, 8(1), 26–42.
- Fox, J. (2015). Editorial, Trends and issues in language assessment in Canada: A consideration of context. Special issue on language assessment in Canada. *Language Assessment Quarterly*, 12(1), 1–9.
- Fox, J., & Artemeva, N. (2011). *Raters as stakeholders: Uptake in the context of diagnostic assessment*. Paper presented at the Language Testing Research Colloquium (LTRC), University of Michigan, Ann Arbor.
- Fox, J., & Haggerty, J. (2014). *Mitigating risk in first-year engineering: Post-admission diagnostic assessment in a Canadian university*. Paper presented at the American Association of Applied Linguistics (AAAL) Conference, Portland.
- Fox, J., Cheng, L., & Zumbo, B. (2014). Do they make a difference? The impact of English language programs on second language (L2) students in Canadian universities. *TESOL Quarterly*, 48(1), 57–85. doi:10.1002/tesq.
- Fox, J., von Randow, J., & Volkov, A. (2016). Identifying students-at-risk through post-entry diagnostic assessment: An Australasian approach takes root in a Canadian university. In V. Aryadoust & J. Fox (Eds.), *Trends in language assessment research and practice: The view from the Middle East and the Pacific Rim* (pp. 266–285). Newcastle upon Tyne: Cambridge Scholars Press.
- Freadman, A. (2002). Uptake. In R. Coe, L. Lingard, & T. Teslenko (Eds.), *The rhetoric and ideology of genre* (pp. 39–53). Cresskill: Hampton Press.
- Gee, J. P. (2011a). *An introduction to discourse analysis: Theory and method* (3rd ed.). London: Routledge.
- Gee, J. P. (2011b). *How to do discourse analysis: A toolkit*. London: Routledge.
- Huhta, A. (2008). Diagnostic and formative assessment. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 469–482). Malden: Blackwell.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

- Leont'ev, A. N. (1981). The problem of activity in psychology. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 37–71). Armonk: Sharp.
- McLeod, M. (2012). *Looking for an ounce of prevention: The potential for diagnostic assessment in academic acculturation*. Unpublished M.A. thesis. Carleton University, Ottawa.
- Office of Institutional Research and Planning. (2014). *Retention and Graduation of Undergraduates for "B. Engineering" – 1998 to 2012*. Ottawa: Carleton University.
- Prior, P. (1994). *Writing/disciplinarity: A sociohistoric account of literate activity in the academy*. Mahwah: Lawrence Erlbaum.
- Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, 7(2), 180–190.
- Read, J. (2012). *Issues in post-entry language assessment in English-medium universities*. Revised version of a plenary address given at the Inaugural Conference of the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ). Australia, University of Sydney. 10 Nov 2012.
- Read, J. (2015). *Assessing English proficiency for university study*. Gordonsville: Palgrave.
- Scanlon, D. L., Rowling, L., & Weber, Z. (2007). "You don't have like an identity: you are just lost in a crowd": Forming a student identity in the first-year transition to university. *Journal of Youth Studies*, 10(2), 223–241.
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.
- Vygotsky, L. S. (1987). Thinking and speech (N. Minick, Trans.). In R. W. Rieber & A. S. Carton (eds.), *The collected works of L. S. Vygotsky: Vol. 1. Problems of general psychology* (pp. 39–285). New York: Plenum Press. (Original work published 1934).