Teaching Mathematics and its Applications: An International Journal of the IMA (2021) **00**, 1–29
https://doi.org/10.1093/teamat/hrab032 Advance Access publication

Performance assessment for mathematics tutoring centres

Carolyn Johns[†,*], Cameron Byerley[‡], Deborah Moore-Russo[§], Brian Rickard[‖], Janet Oien[¶], Linda Burks[#], Carolyn James[†], Melissa Mills[‡], William Heasom[§], Melissa Ferreira[§] and Behailu Mammo[‖]

[†] *Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA*
[‡] *Department of Mathematics, Science, and Social Studies Education, University of Georgia, Athens, GA 30602, USA*
[§] *Department of Mathematics, University of Oklahoma, Norman, OK 73019, USA*
[‖] *Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR 72701, USA*
[¶] *Department of Mathematics, Colorado State University, Fort Collins, CO 80523, USA*
 # Department of Mathematics and Computer Science, Santa Clara University, Santa Clara, CA 95053, USA
† Department of Mathematics, University of Portland, Portland, OR 97203, USA
‡ Department of Mathematics, Oklahoma State University, Stillwater, OK 74078, USA
§ Department of Mathematics and Statistics, Villanova University, Villanova, PA 19085 USA
‖ Department of Mathematics, Hofstra University, Hempstead, NY 11549, USA

[*] *Corresponding author. Email:* johns.125@osu.edu

[Received March 2021; accepted November 2021]

Evaluation has become a common, and even an expected, practice across undergraduate mathematics tutoring centres in the USA, UK and other countries. However, these evaluation efforts could benefit greatly from leveraging organizational theory research. In this study, we situate mathematics tutoring centres as non-profit organizations (NPOs) to consider how an organization performance assessment framework can be adapted to study mathematics tutoring centre data and characteristics. We use qualitative and quantitative data, collected from six mathematics tutoring centres and adapt Lee & Nowell's (2015, *Am. J. Eval.*, 36, 299–319) NPO performance framework to situate our study. Using thematic analysis, the research team underwent iterative cycles of data collection and analysis to code for alignment with Lee and Nowell's framework. By adapting Lee and Nowell's framework to mathematics centres, each of the six centres was given a more relevant lens to consider its performance. Regardless of the university setting, previous evaluations for centres have focused primarily on outputs (e.g., number of visits), behavioural change outcomes (e.g., correlating visits to grades) and client satisfaction outcomes (e.g., student surveys) that ignore the particular context of a centre. However, Lee and Nowell's framework takes into account additional performance indicators that provide a more nuanced understanding of a centre's performance by bringing to light the interplay among its various dimensions. Lee and Nowell's framework allows centres to look beyond outputs and outcomes to understand why these outputs and outcomes come to be. The use of this adapted performance framework, for the six mathematics centres in this study,

allows an interpretation on a variety of dimensions using relevant data while indicating possible areas for change for each centre.

1. Introduction

Mathematics tutoring centres, known as mathematics support centres in the UK and hereafter referred to as mathematics centres or centres, have become a fixture in USA and UK universities (Matthews *et al.*, 2013; Bressoud *et al.*, 2015). Mathematics centres have justified their funding through student feedback, usage data and regression models indicating a correlation between mathematics centre visits and mathematics grades (e.g., Xu *et al.*, 2001; Byerley *et al.*, 2018; Rickard & Mills, 2018). However, the structure of mathematics centres varies across institutions (Byerley *et al.*, 2019; Lawson *et al.*, 2020). Which factors may contribute to mathematics centre effectiveness and how effectiveness should be defined both remain open questions. The overarching aim of our research program is to form a framework that can be used in the evaluation of mathematics centres.

Our research program began as a quantitative study with the intent to measure the effectiveness of mathematics centres by examining the predicted increase in a student's mathematics grade per visit to the centre, accounting for student's incoming ability and other factors (Byerley *et al.*, 2018; Rickard & Mills, 2018). Upon recognizing the variety of centres, we sought to define the characteristics of centre organizational structures in order to understand their similarities and differences (Byerley *et al.*, 2019). Next, we used quantitative and qualitative measures to compare effectiveness across multiple centres (Byerley *et al.*, 2020) and then used these measures to hypothesize on the organizational features of centres that might lead to increased or decreased effectiveness (Byerley *et al.*, 2021).

We recently turned to literature on organization theory to frame our research. In doing so, we recognize that organizations within the same industry can have varying but successful organizational structures based primarily on features of organization size, environment and technology (Child, 1972; Donaldson, 1996). We also came to understand there are several different ways in which organizations evaluate their performance (Akingbola, 2012; Lee & Nowell, 2015). We needed to adjust our original research aim. Rather than using the same definition of effectiveness across various mathematics centres, we needed to tailor the definition of effectiveness to the features of each particular centre.

In this paper, we address the following research question: how can we adapt and utilize Lee & Nowell's (2015) framework to study mathematics centre data and characteristics? Specifically, we

(1) adapt Lee & Nowell's (2015) non-profit organization (NPO) performance framework to mathematics centres and subsequently
(2) apply this adapted performance framework by detailing
    a. how this framework can be used to interpret mathematics centre performance on a variety of dimensions and
    b. how performance on various dimensions is connected, which can be used to explain mathematics centre performance and indicate areas for change.

To answer our research question, we begin by situating mathematics centres as NPOs and examine the prior work and literature which inform our study. Next, we describe Lee & Nowell's (2015) work on performance assessment as it applies to mathematics centres. In our methods section, we describe our mixed-methods case study design. We then discuss our cross-case analysis of six diverse mathematics centres in the USA, using Lee and Nowell's framework to describe and interpret each centre's

performance. We examine the interactions of various measures to explore reasons for performance and avenues for change.

2. Literature Review

2.1 Organization theory

Child (1972) defined a work organization as an organization in which 'work is carried out on a regular basis by paid employees, which have been deliberately established for explicit purposes' (p. 2). Mathematics centres are organizations established for carrying out the work of tutoring students by paid tutors. More specifically, mathematics centres can be classified as NPOs as centres do not exist for the purpose of making a financial profit.

NPOs differ from for-profit and government agencies in multiple ways (Rojas, 2000; Toepler & Anheier, 2004; Akingbola, 2012; Wadongo & Abdel-Kader, 2014). Although NPOs are diverse, they have similarities in terms of their structures and functions. NPOs are established to achieve social objectives and lack a goal of profit maximization (Toepler & Anheier, 2004; Akingbola, 2012; Wadongo & Abdel-Kader, 2014). In the broader university context, mathematics centres may be used to increase profit through student retention, but as a distinct unit, there is no profit mechanism such as student payment per visit. In addition, NPOs have broad and vague missions and complex goals. For example, mathematics centre missions often include phrases such as 'increase student success' and provide an 'active learning atmosphere'. Further, like NPOs, mathematics centres serve a special function of providing services (tutoring) and advocating for a target population (traditionally, at-risk students).

2.1.1 Organizational structure. In our earlier work (Byerley *et al.*, 2019, 2020, 2021), we hypothesized that we could identify structures of tutor centres, which lead to effectiveness. However, upon further review of the organizational theory literature, we found contemporary views of organization theory hold that there is no one best organization structure (Mintzberg, 1980; Donaldson, 1996). Rather, the effectiveness of an organization is impacted by the fit between the organization's environmental factors and the organization's characteristics (Toepler & Anheier, 2004; Akingbola, 2012). Bradshaw (2009) argues, 'what works in one setting, or at one point in time, may not work in another and that efficiency related to ongoing alignment of various [factors]' (p. 62). In other words, there is no one best common structure for all mathematics centres because each centre exists in a different environment. Therefore, rather than using the same definition of effectiveness across various mathematics centres, it is necessary to tailor the definition of effectiveness to the features of each particular centre.

2.1.2 Performance assessment. Due to the key structural and functional differences, performance assessment of NPOs is distinct from performance assessment in for-profit or government agencies. Wadongo and Abdel-Kader (2014) assert, the '"public good" nature of NPOs products and services does not reflect the true market value or price; thus, competition and price cannot be used as a measure of performance' (p. 681). NPOs lack a universal measure of performance such as bottom line used in for-profit organization (Herman & Renz, 2008; Akingbola, 2012). Rather, performance assessment of NPOs utilizes independent, multi-dimensional criteria that include both financial and non-financial measures (Rojas, 2000; Herman & Renz, 2008; Wadongo & Abdel-Kader, 2014; Lee & Nowell, 2015). Like organizational structure, selection of a performance assessment system is contingent on contextual and organizational features such as size, structure and environment (Wadongo & Abdel-Kader, 2014; Lee & Nowell, 2015). Building off of Wadongo and Abdel-Kader, Lee & Nowell (2015) proposed a

seven-dimensional performance assessment framework consisting of (1) inputs, (2) organizational capacity, (3) outputs, (4) behavioural and environmental change outcomes, (5) client satisfaction outcomes, (6) public value accomplishment and (7) network/institutional legitimacy. Lee and Nowell's framework is discussed in-depth as the analytical framework later in this article.

2.2 *Evaluation of mathematics tutoring centres*

Historically, much of the research on tutoring has focused on the practices of tutors and tutor-student interactions (e.g., Arcavi & Schoenfeld, 1992; Graesser *et al.*, 1995; Chi, 1996; Lepper & Woolverton, 2002). Only recently has research examined mathematics centres themselves (Matthews *et al.*, 2013; Mills *et al.*, 2020). A portion of this research involves descriptive reports on the state of mathematics centres at institutions with no evaluation of the centres. Similar to Matthews *et al.* (2013), we focus on mathematics centre research that includes elements of evaluation. In particular, we concentrate on the methods of centre evaluation, as well as gaps in the body of existing research. For more extensive reviews of the literature on mathematics centres, see Matthews *et al.* (2013), Mills *et al.* (2020) and Lawson *et al.* (2020).

Literature on mathematics centre evaluation has primarily focused on evaluation as a means to ensure centre funding. MacGillivray & Croft (2011) argue that while evidence shows a need for mathematics centres, funding for centres is often not secure. Techniques of evaluation generally fall into four categories: centre usage data, links between student grades (or retention) and centre visits, links between student confidence and centre visits and student feedback (Matthews *et al.*, 2013).

Centre usage data is the most commonly used evaluation method reported in a survey of US mathematics centres (Mills *et al.*, 2020) and can be found in a number of papers (e.g., Staddon & Newman, 2003; Dowling & Nolan, 2006; Gill & O'Donoghue, 2007). Usage data, such as number of student visits or number of visits per student, can show student demand for and value of services (Croft, 2009; MacGillivray & Croft, 2011).

Studies examining the correlation between student grades or retention and centre visits are another common means of demonstrating centre value. Several studies at various universities have employed regression models to help account for the biassing self-selection of centre attendance (e.g., Xu *et al.*, 2001; Berry *et al.*, 2015; Byerley *et al.*, 2018; Rickard & Mills, 2018; Rylands & Sherman, 2018; Jacob & Ni Fhloinn, 2019). For example, Xu *et al.* (2001) performed a regression analysis and found that visiting the centre correlated with higher final exam scores in College Algebra when controlling for well-known predictive variables (i.e., gender, SAT, mathematics placement and high school grade point average or, GPA).

Student feedback on satisfaction and confidence is another means of demonstrating the value of centres (Matthews *et al.*, 2013). For example, self-report surveys and interviews have provided evidence that tutoring can lead to improved students' confidence (Croft *et al.*, 2008; Gillard *et al.*, 2011; Parsons *et al.*, 2011; Carroll & Gill, 2012; Wilkins, 2015; Dzator & Dzator, 2020). Centres also seek feedback to understand why at-risk students do not attend tutoring. However, researchers warn that student feedback may be of limited utility due to the overly positive nature of student responses (Booth, 2003; Croft, 2009; Lawson, 2015).

These four types of evaluation are generally only descriptive or comparative between tutoring and no tutoring. While these evaluation methods may demonstrate the impact of mathematics centres in order to obtain or maintain funding, researchers have begun to call for studies which demonstrate strategies which lead to efficiency and effectiveness (Ni Fhloinn, 2009; Kyle, 2010). This type of evaluation can be comparative to an ideal standard, longitudinally within the same organization or to similar organizations

(Herman & Renz, 2008). Currently, there is no established ideal of centre performance. Matthews *et al.* (2013) point out that longitudinal studies can be problematic due to the wide range of other university factors that change year to year. In addition, they argue comparing across institutions is difficult due to reluctance to share data. Only a small portion of the literature has reported on data from multiple institutions (e.g., Gillard *et al.*, 2011; Mac an Bhaird *et al.*, 2011; O'Sullivan, *et al.*, 2014; Byerley *et al.* 2019; Byerley *et al.*, 2020; Mills *et al.*, 2020).

Furthermore, evaluation of mathematics centres has not taken into consideration literature on organizations. As we have argued, mathematics centres are organizations which may be able to leverage the vast body of research on organizations, specifically NPOs. Centre evaluation has elements of organizational performance evaluation practices but has not drawn from that literature. For example, Ni Fhloinn (2009) argues for a multi-faceted evaluation approach that includes correlation of final grades to centre visits, feedback from students and anecdotal information from tutors. Gillard *et al.* (2011) suggest that to show effectiveness, mathematics centres need to demonstrate their impact on the different stakeholders (i.e., the students, the centre and the institution). Matthews *et al.* (2013) also argue the need to align evaluation with the mission of centres. These assertions align with aspects of organizational theory, but no study has utilized or applied organizational theory to mathematics centre evaluation.

3. Analytical Framework

In order to make sense of how mathematics centres might measure their performance, we draw on Lee & Nowell's (2015) framework of NPO performance assessment. Lee and Nowell created an integrated framework utilizing a variety of perspectives from performance assessment literature. They outline seven performance dimensions found in the literature on nonprofit performance assessment: inputs, organizational capacity, outputs, behavioural and environmental outcomes, client satisfaction outcomes, public value accomplishment and network/institutional legitimacy (Table 1). The relationships between these dimensions are illustrated in Fig. 1. Drawing on contingency theory (Mintzberg, 1980; Donaldson, 1996), Lee and Nowell argue contextual factors would lead an NPO to focus on certain performance dimensions. In this section, we provide an overview of Lee and Nowell's framework, highlighting its applicability to mathematics centres (Table 1). We focus on aspects particularly salient to the mathematics centre context and consistent with Lee and Nowell's contingencies for choosing performance assessment practices (Fig. 1).

3.1 Performance dimensions

Lee & Nowell (2015) argue NPOs work under the constraint of budget and resources. For example, mathematics centres are unable to make money to hire additional tutors or improve their physical facilities. *Input* performance examines how resources, such as funding, facilities and equipment, are acquired and used. Common mathematics centre resources include funding for tutors' salaries, the physical space provided by the university to the centre and the number of hours that staff (or faculty) serve as centre leaders. In the mathematics centre context, the strength of the relationship between the centre and its resource providers (e.g., department chairs or upper administration) is a measure of resource acquisition.

Organization capacity measures the efficiency and effectiveness of the organization's processes towards furthering their mission, as well as the organization's ability to innovate and adapt (Lee & Nowell, 2015). The organizational capacity includes centre processes such as how students find

TABLE 1. *Performance measures for centres (adapted from Lee & Nowell, 2015, p. 302)*

Performance dimension	Definition	Examples of performance measures in mathematics centres
Inputs	Acquirement and use of resources	• Funding for tutors • Physical space (location, square footage) • Ability to acquire and manage tutors • Contracted hours for non-salaried staff • Relationship with funders
Organizational capacity	Effectiveness of processes to use resources efficiently; capacity to innovate	• Employee satisfaction, motivation, retention and capabilities • Employee training • Critical operating processes such as marketing the centre or the check-in process • Capacity of the centre to adapt to changing needs
Outputs	Countable goods and services	• Number of students served • Number of courses served • Number of hours open • Number of services provided (e.g., tutoring, review sessions, calculator and textbook loans) • Physical space (e.g., layout, décor, signage)
Outcomes: behavioural and environmental changes	Intended change in target population	Impact on mathematics course grades • Impact on students' confidence or feelings of belongingness • Impact on students' use of self-regulatory processes or other study resources
Outcomes: client satisfaction	Satisfaction of customers; meeting needs of customers	• Percent of eligible students using services • Average visits per student attending mathematics centre • Student satisfaction surveys
Public value accomplishment Network/institutional legitimacy	Impact on community/society Relationships with stakeholders and industry community; adherence to mission	Community's perceived value of the centre's impact • Satisfaction of administrators and instructors • Partnership/collaboration with internal and external tutor centres • Coherence of activities with stated mission

information on tutoring; how students register for services; and how tutors are hired, trained, motivated and retained. Centres need to be able to innovate by altering tutor operations to meet changing needs or streamlining processes to lower costs. The time available to centre leaders to effectively select and implement changes is one way to measure capacity for innovation.

Outputs are 'countable goods and services obtained by means of nonprofit activities and direct products of activities for achieving the mission' (Lee & Nowell, 2015, p. 306). For example, outputs include the number of student visits, the number of hours the centre is open, the number of courses tutored and the services the centre offers. Outputs can be compared with inputs to examine centre efficiency. Comparative ratios of centre measures such as tutor cost per student visit (i.e., total tutor wages divided by number of student visits in a year) and utilization of facility space (i.e., total centre square feet times number of hours open divided by number of student visits in a year) assess input to output efficiency.

Outcomes are separated into two dimensions: behavioural and environmental changes, as well as client satisfaction (Lee & Nowell, 2015). *Behavioural and environmental changes* examine the 'state of the target population or social condition the program is supposed to have changed' (Rossi *et al.*, 2004, p. 204). Lee and Nowell argue this dimension extends beyond outputs to consider the impact on the target population. For mathematics centres, this includes examining the increased skills, knowledge

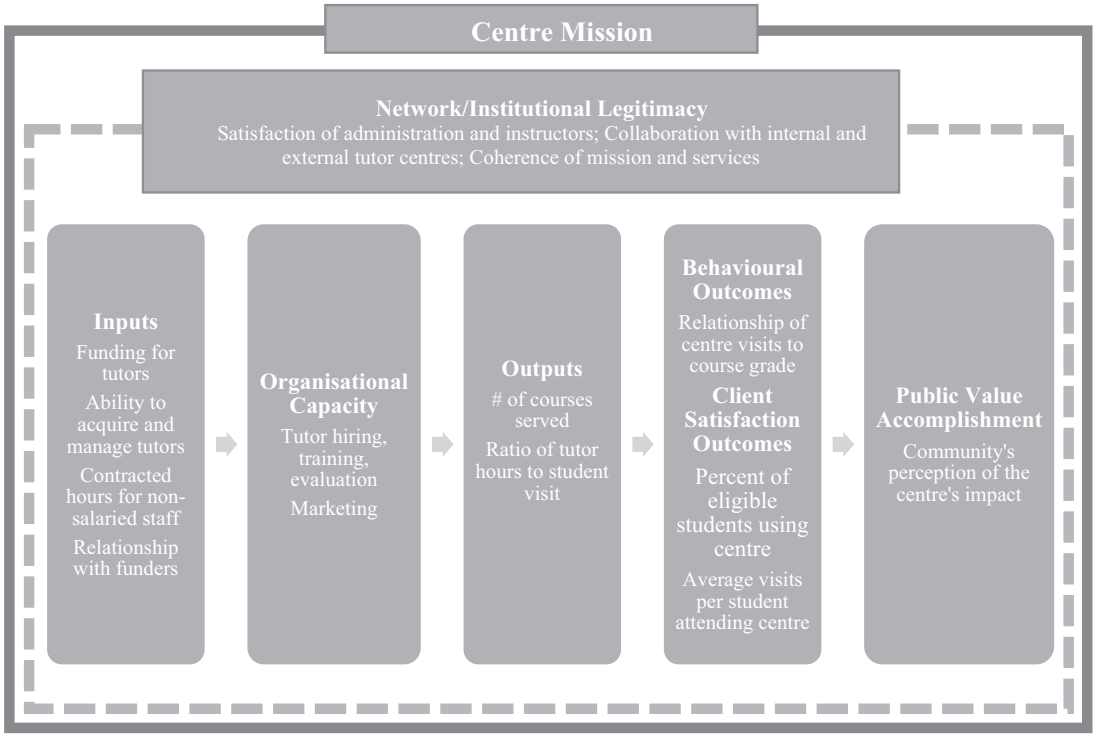

FIG. 1 Relationship between performance dimensions for mathematics centres. Each performance dimension should be considered as it pertains to the advancement of the centre's mission. Adapted from Lee & Nowell's (2015, p. 305).

or learning gained by the students from visiting the mathematics centre. Assessments of the extent to which students' confidence or feelings of belongingness have been altered, or the changes in students' use of self-regulatory processes and other study resources are also measures of behavioural and environmental changes. The second outcome dimension, *client satisfaction*, measures the 'extent to which the organization satisfied and met the needs of the [target] population' (Lee & Nowell, 2015, p. 303). In mathematics centres, this is often measured by the percent of eligible students using their services, the number of visits per student using the centre and student satisfaction surveys.

Public value accomplishment is the 'ultimate value/impact the organization hopes to create for the community/society' (Lee & Nowell, 2015, p. 303). In other words, it is the extent to which the organization is achieving the social ambition outlined in its mission statement according to outsiders' evaluations and opinions.

Network/institutional legitimacy is 'how an organization has managed its relations with other stakeholders and established a reputation for trustworthiness and excellence within this broader network' (Lee & Nowell, 2015, p. 308). For example, it involves assessing how well the mathematics centre works with other on-campus and off-campus tutor centre organizations with similar goals. It also includes the relationships that math centre staff have with individuals who can supply other resources the centre might use. The satisfaction of the funders and stakeholders, typically department faculty and administration,

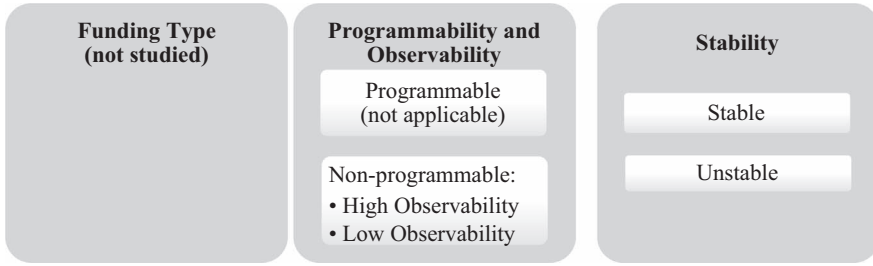

FIG. 2 Lee & Nowell's (2015) three, independent environmental contingency factors.

is also part of network. Institutional legitimacy includes the extent to which the centre's activities match the stated mission.

3.2 *Contingencies*

While all of the performance dimensions (Table 1) impact each other, organizations should focus on particular dimensions based on environmental contingencies (Lee & Nowell, 2015). Lee and Nowell suggested three contingency factors: (1) funding type, (2) the extent to which the work is programmable and observable and (3) the stability of the organization's environment (Fig. 2). Lee and Nowell discuss each of these contingencies individually and do not comment on how different combinations of contingencies interact. In addition, Lee and Nowell only comment on how instability impacts performance practices but do not elaborate on stability. In our discussion and analysis, we detail how these contingencies inform performance assessment practices. At present, we discuss how these contingencies can be applied to mathematics centres.

Due to the data available, we do not consider the funding type contingency in this study. With respect to programmability, the work of tutoring is non-programmable as it cannot be carried out via predetermined procedures and protocols (Eisenhardt, 1989). So, only non-programmability needs to be considered in light of the degree of observability (Lee & Nowell, 2015). Mathematics centres have high observability when centre leaders are able to directly observe tutoring and low observability when constraints on centre leader time, the physical space of the centre or the private nature of one-on-one appointments impede a centre leader's attention to the service the centre provides.

Lee & Nowell (2015) consider the stability of the environment as the third contingency factor. Stability is impacted by operating decisions at both the department and centre levels; instability in mathematics centres is caused by frequent employee turnover or factors that impact a tutor's ability to be prepared mathematically. Centres in departments where courses are uniform due to coordination of topic coverage, scheduling and assessments have more stable environments. In contrast, centre tutors working for departments without course coordination might experience a wide variety in topics covered, textbooks used, assessments given and schedules for the same course taught by different instructors. Tutors who specialize in tutoring only one or a few related courses have a more stable environment than tutors in centres who serve as generalists working with students in a wide variety of courses. Lack of course coordination and specialized tutoring provide tutors a less stable environment due to constant changes in the mathematics content they must address in their tutoring.

TABLE 2. *Mathematics centre groupings for case selection*

Grouping ∗Selected case	Mathematics centre	Size (# of eligible students)	Observability (supervisor or lead tutor present)	Stability (course coordination)
Group A	Cat	Large	High	High
Group A∗	Gorilla	Large	High	High
Group A	Whale	Large	High	High
Group B∗	Dog	Small	High	High
Group C∗	Hamster	Large	Low	High
Group C	Fish	Large	Low	High
Group D∗	Bird	Small	Low	High
Group E∗	Horse	Small	High	Low
Group F	Goat	Small	Low	Low
Group F∗	Dolphin	Small	Low	Low

4. Method

This study was conducted as a secondary analysis of previously collected data (Smith, 2008) with some additional data collected specifically for this study. We employed a mixed-methods, multiple case study design (Yin, 2018). Each unit of study was an organization, an individual mathematics centre. Our cases were designed to be explanatory, heuristic case studies (Mitchell, 1983), in order to build theory on how Lee & Nowell's (2015) performance assessment framework can be adapted for mathematics centres and used to interpret centre performance. We examine how performance dimensions impact each other, which can be used to explain centre performance and indicate areas for change. We utilize both qualitative and quantitative data to account for each centre's performance on various measures from Table 1.

4.1 Case selection

This study was conducted as a secondary data analysis of the data presented in Byerley et al.'s (2019, 2020, 2021) work. The 10 cases in the original data were collected as a convenience sample (Byerley *et al.*, 2020). They represented centres whose leaders were involved in the Research in Undergraduate Mathematics Education mathematics centre community, interested in participating in the study and able to collect the necessary quantitative data.

From the centres used in Byerley *et al.*'s (2020) study, we chose various centre qualities to form groups (Glaser & Strauss, 1968) to be studied. As environmental factors impact both appropriate organizational structures (Donaldson, 1996) and performance assessment (Lee & Nowell, 2015), we grouped centres by size, observability and stability.

We operationalize size of the centre as the number of students eligible to use the services and split the size of the institutions into two even groups: small and large. Observability was operationalized by whether centre leadership was present with the tutors at least 50% of the time and split the institutions into two even groups with high or low observability. While there are other contributing environmental factors, we operationalized stability as the presence of course coordination. Centres where less than 50% of the students being tutored came from coordinated courses were considered to have low stability.

These three factors produced a possible eight groups of the various combinations of the three contingency factors. However, the 10 centres represented only 6 of the possible groupings (Table 2). In particular, no large schools had environmental instability.

Based on our groups, we selected mathematics centres as typical cases (Thomas & Myers, 2015) in order to maximize differences across the centres by size, observability and stability (Glaser & Strauss, 1968). Thus, although the original 10 cases were selected based on convenience, the case studies presented here were chosen purposefully based on their ability to develop theory. When a group contained more than one centre, the centre with the more extensive qualitative data from the original data set was selected in order to include as many points of data as possible. From the six groups, we selected the Gorilla, Dog, Hamster, Bird, Horse and Dolphin mathematics centres for study.

4.2 Data collection

During the original collection of the secondary data, all centres were asked to fill out an open-ended survey based on Byerley *et al.*'s (2019) characteristics of tutor organizations and to submit quantitative data on their centre. The data included the results of a multiple regression model to capture the correlation of centre visits to final course grade (see Byerley *et al.*, 2020, 2021). Qualitative descriptions include the relationship of the centre to mathematics department faculty, the percentage of time the centre leader was expected to spend operating the centre, the tutor training, the operation hours, the services provided and the physical space. Quantitative data included the number of students eligible to use the centre (counted as number of students enrolled in courses served by the centre), the number of student visits, the square footage of the centre and the tutor hours (taken as the sum of all hours worked by all tutors in a given semester). Quantitative metrics were calculated from these data to determine the percentage of those students who visited the centre, the average number of visits to the centre per visiting student, the tutor hours per student visit and the square footage per visiting student.

For the regression analysis, it was deemed desirable to have a standardized analysis across centres in order to compare results. Based on data available to participating centres, four factors were included in the regression analysis: (1) student visits to the mathematics centre, (2) high school grade point average, (3) standardized entrance test scores and (4) course letter grades converted to grade points.

Depending on the year data were available, quantitative data from each institution were collected for the fall semester of 2017 and/or 2018. Only students enrolled in mathematics courses the centre purposefully served were included in data analysis. Students with missing data and students who withdrew from the course were removed from the regression analyses but were included in the quantitative metrics such as eligible and visiting students. Students enrolled in multiple mathematics courses were treated as separate data points, with the number of visits to the centre split equally between enrolled courses.

For this study, the researchers collected additional qualitative data from the six selected universities. Centre leaders were emailed a qualitative survey of open-ended questions based on Lee & Nowell's (2015) framework to fill gaps in the initial data. The individual who collected data from the primary study emailed these questions to participants, blinded the results and sent them to the first author of this study. After a second wave of coding, answers to remaining follow-up questions were collected.

4.3 Data analysis

In the primary study, multiple regression analyses were conducted on the data from each mathematics centre using course grades as the dependent variable and student visits to the centre, high school grade point average and standardized test scores as the independent variables. Analyses combined students from all centre-eligible courses within each respective university.

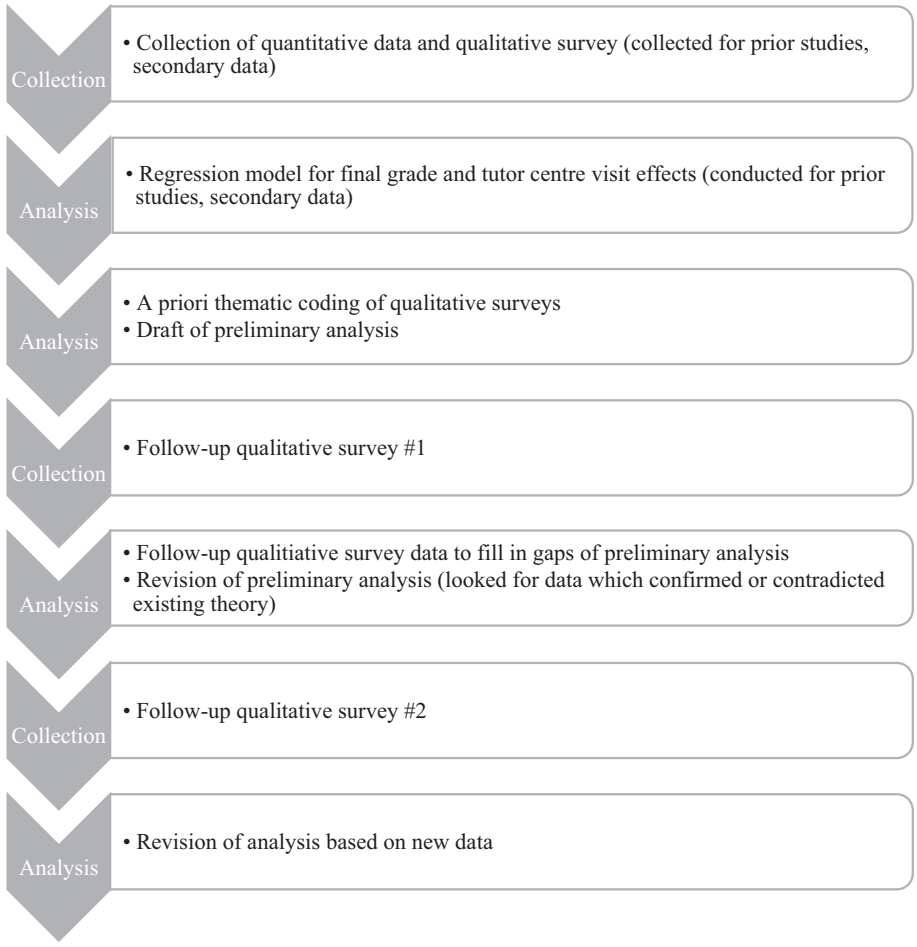

FIG. 3 Iterative data collection and analysis.

For this study, analysis of the data was an iterative process (Fig. 3) using constant comparative methods to look for patterns and themes (Charmaz, 2014). We used deductive coding (Braun & Clarke, 2006) on each centre's qualitative survey and quantitative data, using Lee and Nowell's eight performance dimensions as codes. We then wrote explanatory case studies (Yin, 2018) regarding each centre's performance on measures from Table 1. For analysis, we grouped the six cases into two larger grouping of low and high observability, based on Lee & Nowell's (2015) contingencies. We compared cases within and across these two groupings. With each wave of data collection and analysis, the theory was refined. Respondent validation was then employed (King & Brooks, 2018) to ensure validity of the analysis.

5. Results and Discussion

To interpret each centre's performance, we align ourselves with Herman and Renz's (2008) assertion that NPO effectiveness is comparative. Organizations may be compared with themselves at an earlier

TABLE 3. *Focal performance dimensions based on contingencies*

Contingency	Focal performance dimensions	Centres
High observability	• End-to-end process of inputs, capacity and outputs • Behavioural outcomes	Gorilla, Dog, Horse
Low observability	• Capacity • Client satisfaction outcomes	Bird, Dolphin, Hamster
Instability	• Network and institutional legitimacy	Horse, Dolphin

point in time, to similar organizations or to an ideal. Within our analysis, we compare across centres and a hypothetical ideal. Herman and Renz also argue 'there is no commonly agreed basis for judging [nonprofit organisation] effectiveness' (p. 404). They assert effectiveness is a social construct as the measures mean nothing until someone interprets them. We acknowledge the results may be interpreted differently by others; however, as practitioners with many years of experience working in centres, we feel we have the knowledge to provide plausible interpretations of the measures.

To examine centres' performance on Lee & Nowell's (2015) performance dimensions (Table 3), we examine the centres in two groups based on their observability contingency. We first examine centres with high observability where a supervisor is physically present and observing the tutors at least 50% of the time (Gorilla, Dog, Horse). For these centres, Lee and Nowell propose focusing on end-to-end process of inputs, capacity and outputs, as well as their behavioural outcomes. We then examine centres with low observability (Bird, Dolphin, Hamster), focusing on their organizational capacity and client service outcomes. We also note that Horse's and Dolphin's instability, resulting from less than 50% of courses served having coordination, can lead to a performance assessment that focuses on network and institutional legitimacy. Within these discussions, we also explore how each centre's relationships with its funders and the perceived value of the centres may influence centres' abilities to increase or maintain funding. While we do not separate centres by Lee and Nowell's funding contingency, for any funding model, the relationships with funders or the public's perceived value of the centre could be considered.

5.1 Background and terminology

To help orient the reader, we introduce common terminology we will use throughout the discussion and results (Table 4).

The six centres studied differed in other regards besides the contingencies. Table 5 provides additional background information for each centre.

5.2 Centres with high observability

As centres with high observability of tutors (i.e., a supervisor is physically present and observing the tutors at least 50% of the time), Gorilla, Dog and Horse should focus on the end-to-end process of inputs, capacity, outputs and behavioural outcomes (Lee & Nowell, 2015) when compared with professional standards (Table 3) (Lee & Nowell, 2015). As there are no established professional standards for mathematics centres, we compare across institutions to examine centre performance relative to other centres. With respect to professional standards of tutors and tutoring, we rely on literature on effective tutoring and teaching.

TABLE 4. *Definitions for mathematics centre terminology*

Terminology	Definition
Mathematics centre or centre	Centre that exists for the purpose of helping students in mathematics (and often statistics) primarily through tutoring
Tutoring centre	Other centres on campus which tutor non-mathematics courses
Academic support group	Centralized group of multiple tutor centres on campus
Centre leaders or leadership	All mathematics centre employees with a rank above tutor
Director	Full-time employee who oversees the mathematics centre
Director's supervisor	Individual to whom the director reports
Assistant director	Full-time employee who assists the director in management of mathematics centre
GA	Graduate students who tutor, as well as instructors (with graduate degrees or who are graduate students) who both tutor and teach
Student managers	Student employees who assist in management of mathematics centre and who rank above the tutors
Course coordinator	Individual tasked with creating common syllabus, calendar, homework and exams across all sections of a course

TABLE 5. *Background information on mathematics centres*

Institution	# of courses tutored	Type of tutors	Centre leadership	Director's supervisor	Housing unit
Bird	11	Undergraduate	Director	Director of academic support group	Academic support group
Dog	5	Undergraduate; GAs	Director, two assistant directors	Chair of Mathematics Department	Mathematics Department
Dolphin	7	Undergraduate	Director	Chair of Mathematics Department	Mathematics Department
Gorilla	12	Undergraduate, Optional for GAs	Director, student managers	Chair of Mathematics Department	Mathematics Department
Hamster	10	GAs	Two co-directors	Chair of Mathematics Department	Mathematics Department
Horse	13, any math question	Undergraduate, GAs	Director	Associate vice provost	Academic support group

As can be noted in Fig. 1, there is a flow of impact from inputs to capacity to outputs to behaviour. We begin our discussion at the start of this flow by examining centres' performance on inputs to capacity. We then explore performance in the behavioural outcome dimension, noting behavioural outcome measures that align with centres' mission statements. Within this discussion, we examine performance on the inputs, capacity and outputs chain that flows into behavioural outcomes. We also examine Horse's network and institutional legitimacy dimension given Horse's instability (Table 3). Finally, we illustrate how each centre's aforementioned performance, relationship with funders and public value might be leveraged for additional funding. In Table 6, we offer a summary of each centre's mission to frame the objective to which the centre should be assessed.

5.2.1 Inputs to organizational capacity. In this section, we examine each centre's inputs, organizational capacity (processes) and the relationship between the two. For inputs, we focus on the amount of time centre leadership has for managing the centre. We also include collaborative network ties, which can help reduce the leadership's workload, allowing leadership to 'off load' some of the intellectual work

TABLE 6. *Centre mission or Director's interpretation of mission when no official statement exists*

Centre	Mission statement
Bird	To provide support for active learning.
Dog	The mission of the centre is to improve student learning and academic performances for all students in our calculus courses. We support instruction that carefully considers student thinking and engages students as active participants in their learning process.
Dolphin	The centre is a community of mathematical thinkers who connect with each other in order to learn together and achieve their goals. Students learn through individual tutoring and study groups. Tutors and students also have some opportunities to grow and work together during special events throughout the quarter.
Gorilla	We provide a place for students to work on their homework and receive help from tutors if needed. We hope to create a place for students to belong and to be productive on their mathematics homework and to assist them in succeeding in their mathematics courses.
Hamster	The mission of the centre is to emphasize relevance, cultivate understanding and promote utilization of mathematics and statistics by providing an engaging and collaborative educational environment for our diverse community of learners.
Horse	We create engaging and effective learning experiences for students studying mathematics and statistics. Engaging means: friendly and approachable tutors, interactive sessions and user-friendly technology. Effective means: building skills and confidence, real-time learning feedback and progress in the classroom.

by pulling ideas from similar organizations and asking others to collaborate. We examine how inputs of centre leadership time and ability to hire/manage tutors can impact processes of employee hiring, training and evaluation.

Inputs. The input of centre leadership time varied at each institution. Horse and Gorilla each had a single director who spent 67–75% of their time on tutoring-related activities, while Dog employed a director and two assistant directors. Gorilla and Horse both utilized their tutors to extend oversight of the centre through the use of student managers. Horse's student managers oversaw progress of tutoring sessions in-the-moment, while Gorilla's student managers were more involved. Each student manager in Gorilla had a small group of tutors they were responsible for managing, observing and providing feedback. Gorilla's director held weekly meetings with the student managers to discuss any issues that came up in the week and plan future training or fun activities. The Gorilla student managers helped identify tutors doing exceptionally well and helped select the tutor of the month. They also helped with hiring, initial scheduling, coordinating shift covers, scheduling changes and updating policies and procedures each semester. Student managers allowed Gorilla's director to focus on higher-level aspects of the centre by off-loading the day-to-day administrative tasks to them. Through Gorilla's use of student managers and Dog's extensive leadership team, both centres allot a substantial amount time available to hiring, training and evaluation, as well as having time to consider innovations.

In addition, each centre was able to be more efficient with their time through collaboration. Horse's position as a sub-centre of the Academic Support Services extended their network to other on-campus tutor centres. In addition, Gorilla's director and Dog's director and one of Dog's assistant directors were given time to dedicate to research on mathematics centres which led to collaboration with mathematics centre leadership at other institutions. By forming research partnerships with centre leaders at other institutions, Gorilla and Dog were able to draw on those resources for innovation. In addition, through conducting research on mathematics tutoring, Gorilla's and Dog's leaders were able to draw on best-practice findings from existing literature, as well as findings from their own research. While Horse did not have the time for external relationships at the time of the study, through the addition of a part-time assistant director, the leadership team now has additional time for collaboration and innovation.

The amount of centre leadership time can provide varying constraints or affordances for enacting organizational capacity processes of tutor hiring, training and evaluation.

Organizational capacity. The scope of a centre's mission dictated appropriate employee hiring, training and evaluation processes. Horse endeavoured to answer any mathematics or statistics questions from any student and, thus, needed tutors who could answer a wide range of questions which led to a focus on tutor's mathematical knowledge. On the other hand, Gorilla's and Dog's missions targeted fewer courses which allowed them to hire and train tutors based on ability to assist students on a few key courses. In addition, each centres' input of ability to hire and manage tutors impacted their processes.

Gorilla, Dog and Horse all directly hired undergraduate tutors; however, Dog and Horse also had graduated assistants (GAs) who were required to work in the centre. Dog and Horse did not have any input with regard to the hiring or assignment of GAs. This inability to directly hire GAs meant they could not seek out specific desirable characteristics which affect tutor/student interactions and thus behavioural outcomes.

Undergraduate tutor hiring practices varied among centres. Horse focused on the number of mathematics courses taken and grades received with availability, 'eagerness' and diversity considered second. For Gorilla and Dog, while grades were important, they were not the primary factor in hiring decisions. Gorilla's director 'value[ed] personality and trainability over high grades'. Dog's assistant director was not 'just looking for someone with good grades, but someone who tried to understand the concepts of the course, asked questions, participated and got along with their peers'. The centres' differences in qualifications were reflected in their hiring practices. Gorilla, Dog and Horse required tutors to submit a list of mathematics courses taken and grades earned. Gorilla and Dog also required applicants to provide an instructor reference. References were contacted to not only comment on students' mathematical performance in the class but also their participation and communication skills including how they got along with peers. Applicants at all three centres were interviewed by a centre leader; however, the focus on the interviews differed. Due to Horse's need for tutors who could answer a wide range of questions, the director prioritized mathematics knowledge in reviewing applications but during the interview process, the director focused on enthusiasm and people skills. On the other hand, because Gorilla's tutors only needed to be familiar with content from a few courses, Gorilla was able to consider applicants based on other criteria. The affordance combined with Gorilla's director's time to conduct research and network led Gorilla to implement practices which were grounded in research. Gorilla's hiring interviews focused on tutors' trainability on pedagogical techniques as demonstrated by performance on a mock-tutoring scenario. Dog conducted hiring interviews similarly. Research has shown that good grades in mathematics courses are not enough to be an effective teacher (Ball *et al.*, 2005). Teachers also need to have productive dispositions and beliefs to support development and enactive of productive mathematics teaching processes (Thompson, 1992). These dispositions and beliefs are difficult to change (Cooney *et al.*, 1998); so, targeting tutors already holding productive dispositions may lead to greater uptake in desired tutoring practices.

Training also varied between the three centres. Drawing on their research knowledge, Gorilla and Dog aimed to develop tutors with well-rounded mathematical knowledge for teaching (Ball *et al.*, 2008), consisting of both content and mathematics specific pedagogical knowledge. In addition, with tutors serving a narrow range of students, training was able to target specific knowledge and skills required for a few courses. Horse on the other hand had to prepare their tutors to answer any mathematics or statistics questions that were brought to the centre, as well as questions on statistical software. As such, Horse focused on content and general pedagogical skills. At Horse, undergraduates received the same tutor training as GAs, focusing on centre procedures with optional training on specialized course content and general pedagogy topics such as student disabilities and handling difficult situations. However, since data

collection, Horse's assistant director has given leadership time to innovate and expand training. At Dog, training was different for GAs and undergraduates; however, both experienced content and pedagogical training. Undergraduates received tutor-specific pedagogical training conducted in partnership with the academic support group on campus, while mathematics content and mathematics pedagogy specific training occurred within the mathematics centre. Both GAs and undergraduate tutors received continuous content training via attending lecture and weekly course led by the course coordinator. The course coordinator would go over upcoming content, focusing on productive conceptual understandings of the material, ways of student thinking and teaching techniques. Further, because GAs' training for teaching was also conducted by centre leadership, it reflected the active-learning pedagogy encouraged in tutoring. At Gorilla, training also entailed a mix of content and pedagogy. Training at Gorilla included 8 h per semester which tended to 'involve a lot of role play and analysis of tutoring transcripts'. Gorilla's director sought to develop these skills through training akin to video clubs (van Es & Sherin, 2008, 2010; Santagata & Guarino, 2011; Santagata & Yeh, 2014; Karsenty & Sherin, 2017; Schoenfeld, 2017) with an aim to produce reflective practitioners (Schon, 1987; Karsenty & Sherin, 2017). Video clubs have been shown to improve teacher's pedagogical skills (van Es & Sherin, 2008, 2010; Santagata & Yeh, 2014; Schoenfeld, 2017). In addition, reflective practitioners continuously look to improve their practice (Schon, 1987; Karsenty & Sherin, 2017).

Evaluation of undergraduate tutors was similar at all three centres. Due to the frequent presence of a centre leader in the centre, most evaluation and feedback to undergraduates occurred informally and in-the-moment. When tutors were not meeting expectations or received negative feedback, leadership would meet with the tutor. Horse and Gorilla also had more formal feedback mechanisms. Both Horse and Gorilla collected student feedback that was shared with tutors. Gorilla also collected one tutoring observation for each tutor and then met with each tutor to give personalized feedback. With regard to GA evaluation, Dog and Horse were also reliant on the broader university/departmental structure for the centre leadership's role in GA evaluation. At Dog, the departmental structure and culture of active-learning pedagogy allowed the centre leaders to report performance concerns with GAs to the course coordinators or graduate committee. At Horse, the director was the supervisor of the GAs working in the centre and could directly discuss performance issues or take away the student's assistantship.

We have described how each centre's mission and inputs of centre leadership time and the ability to hire/manage tutors impacted organizational capacity of tutor hiring, training and evaluation. Each centre's inputs and organizational capacity will subsequently impact outputs and behavioural outcomes (Fig. 1). Next, we continue along the chain of performance dimensions to examine outputs and behavioural outcomes.

5.2.2 Outputs to behavioural outcomes. In this section, we examine each centre's outputs, behavioural outcomes and the impact of the centres' missions on outputs and selection of behavioural outcome measures. For outputs, we focus on the input/output ratio of tutor hours per visit and number of courses served. We then discuss various behavioural outcome measures that would align with each centre's mission but report only on the quantitative regression model.

Outputs. We begin by examining the outputs dimension using measures of the input/output ratio of tutor hours per visit and number of courses served. To examine each centre's efficient use of tutor funding, the tutor hours (input) was divided by the number of student visits (output). This ratio was then converted to average time in minutes a tutor could spend per student visit for each of interpretation (Fig. 4). Compared with other institutions, on average, the time tutors at Gorilla and Dog could spend with a student was lower than other centres, while at Horse, this measure was higher than other centres. A higher-than-average measure compared with other centres would indicate they have sufficient resources

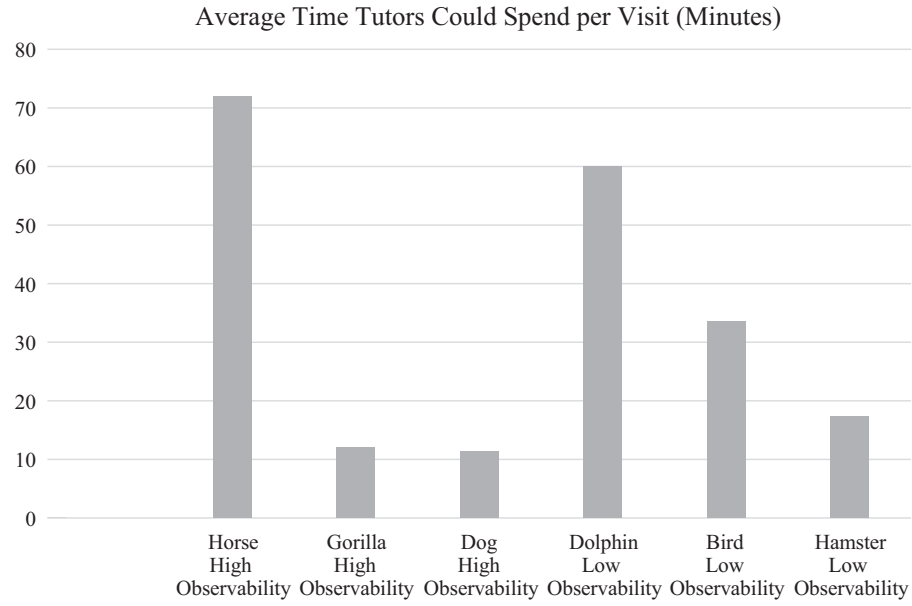

FIG. 4 Input–output ratio of tutor hours/visit for centres as sorted by observability.

to serve students while a higher-than-average ratio may indicate inefficient use of resources. The number of courses served (Table 5), as dictated by the centre's mission (Table 6), varies across the three centres. Dog's mission was to serve calculus students and served only five courses. Gorilla's mission entailed assisting students in first- and second-year courses, generally through differential equations and served 12 courses. Horse, however, had a broad mission to assist with any mathematics question. Officially, they served 13 courses; however, any student could use the centre. As previously noted, the differences in missions impacted tutor hiring, training and evaluation. We now explore how these differences in varying outputs, in turn impacts behavioural outcomes.

Behavioural outcomes. When selecting behavioural outcome measures, it is important that they align with the centre mission. Gorilla and Horse both mention affective goals of belongingness and confidence which should lend to positive student responses to mathematics via an attitudinal survey. To measure student learning and skill building beyond course grades, a conceptual pre/post-test could be employed. In addition, all three centres mention success in mathematics courses as an aspect of their mission. While this can be measured in many ways, for this study, we used a regression model to examine the relationship between centre usage and course grade. Based on the model, after controlling high school GPA in mathematics courses and standardized test scores, for each visit to the centre, students' final grades are predicted to increase 0.015 grade points at Gorilla ($F(3, 2733) = 209.1$, $P < 0.001$, $R^2 = 0.186$) and 0.035 grade points at Dog ($F(3,1000) = 115.5$, $P < 0.001$, $R^2 = 0.257$). While Horse's quantitative model did not indicate a significant effect of visits on mathematics course final grades, this is only one of many possible ways to measure the behavioural outcome of student success. In particular, due to the small number of variables included in the model, centres with access to additional relevant predictors may find different results after controlling for other variables. In addition, Horse's model includes only students in the 13 targeted courses and as such does not include students who received

help from the centre outside of those courses. In addition, as inputs, organizational capacity and outputs flow into behavioural outcomes, Gorilla and Dog may have each benefited from the centres' higher amounts of time leadership dedicated to their centres and a focus on fewer courses which allowed for more extensive and targeted training.

5.2.3 Horse's instability: Network and institutional legitimacy. Horse's performance assessment should also take into consideration its instability (Table 3). Lee & Nowell (2015) do not discuss how the overlap of instability and high observability impacts performance assessment. It may be the case that instability overshadows observability. If the environment is constantly changing, appropriate tutor interactions are also in constant flux creating difficulties with supervision (observability). Not only Horse lacked stability based on lack of course coordination, but also the centre served a wide range of students aiming to answer any mathematics question. Due to Horse's instability, network and institutional legitimacy should be considered as a more appropriate performance assessment.

In examining Horse's network, Horse's leadership was well connected to other tutor centres on campus. Horse's director was supervised by the associate vice provost, who held monthly meetings with all academic support group leaders on campus. During these meetings, tutor centre leaders would share updates on any changes, as well as usage data. Horse's Director used this network to gather ideas for changes to the mathematics centre and market the centre's services, which can help advance the centre mission. With regard to institutional legitimacy, it may be beneficial for Horse's director to consider whether its services align with their mission. Horse's director reported that while the centre was required to serve students in the calculus and statistics sequences, serving students outside of the sequences and fielding mathematics/statistics questions from courses outside of the mathematics department were optional services provided by the Horse.

5.2.4 Relationship with funders. We now explore how centres' relationship with funders, an input, can be paired with the centre's public value and performance assessment to justify increased funding or build support for change in centre practices. Performance assessments are often used as leverage for funding (Carmen, 2011), and both Dog and Horse had been successful in using data to support additional funding in the past.

Centres with strong ties to their funders will likely have greater success in gaining additional funding. Gorilla and Dog both had strong relationships with administration. At both institutions, at least one centre leader reported directly to the department chair. In addition, Dog's leadership had administrative connections to the dean and provost, who were influential in obtaining funding for the centre, and took an interest in centre activities. At Dog, the department chair participated in tutoring events, the dean would walk by the centre time to check on its use, and the provost was previously a professor in the mathematics department. These relationships gave leadership at Gorilla and Dog direct lines to funding decision-makers.

Leadership at Gorilla and Dog recognized they did not have sufficient resources as they stated that given additional funding, they would spend it on acquiring additional tutor hours. To justify additional funding, Gorilla and Dog can present to funders a performance assessment including the previous analysis to demonstrate funding for tutors is the primary obstacle for improved centre performance and achieving each centre's mission. For example, Dog's assistant director stated that, 'during crowded times it would be inaccurate to say that our tutoring was "student-centred" and "engages students as active participants in their learning process".' When busy, the tutors often worked through the problem on the board with a small group and, instead of helping students understand a concept, the tutor would move on to the next in line. However, 'when less busy some tutors definitely provide student-centred instruction.'

Providing evidence that the centre's tutors enacted active learning techniques when the centre was less busy but struggled to meet that mission objective when busy could help to demonstrate it was not due to training or other factors that tutors failed to implement active learning, but rather time per student, which could be addressed by additional funding. While the behavioural outcome performance can be used as evidence that Dog and Gorilla are fulfilling their missions, it is not known to what extent the predicted increase in grade points per visit would change if students had more time with a tutor per visit. Dog might also use evidence of the relationship of tutor behaviours to the number of students in the centre to argue for additional funding.

Gorilla and Dog have strong ties to instructors in the department who can voice their perceived value of the centre to administrators. At Dog, GAs were required to hold office hours in the centre; at Gorilla, this was optional. In addition, leadership at both centres taught courses in the mathematics department and worked closely with course coordinators. At Gorilla, the Director had weekly contact with course coordinators who also occasionally facilitated tutor trainings. At Dog, the centre leaders themselves were coordinators, and tutors attended lectures and course instructor meetings. These close ties with instructors added to each centre's perceived value within the department and integrated the centre into the mathematics department network.

At Horse, leadership may take into consideration whether to reduce their services to only the courses they are required to serve. Removing these optional services may allow Horse to focus on its required mission, creating less instability for tutors and improving tutors' ability to help targeted students. However, before making change, Horse's director should take into account the provost and other stakeholders' perceived value of the centre. If the perceived value is the centre's ability to help anyone with a mathematics or statistics question it would be ill-advised to narrow the scope of tutoring. Horse's strong relationship with the provost's office puts them in a good position to have these conversations.

5.3 Centres with low observability

Client satisfaction outcomes, organizational capacity and institutional legitimacy are focal performance dimensions for centres with low observability (Hamster, Bird, Dolphin). Due to the inability to observe tutors' work, centres with low observability must rely on client satisfaction outcomes to evaluate if the centre is meeting its' mission. Client service outcomes include quantitative measures such as percent of eligible students using services and average visits per student attending, as well as qualitative client feedback. While quantitative measures are easier to collect, qualitative feedback can explain quantitative results. These client service outcomes can point to areas within organizational capacity (such as marketing, tutor hiring/training) which are positively or negatively impacting client satisfaction (Fig. 1). We discuss how changes to organizational capacity can improve client satisfaction outcomes within the analysis. We split our analysis into two sections. The first section examines client satisfaction as it relates to 'getting students in the door' as measured by the percent of eligible students attending. In the second section, we examine client satisfaction in terms of retention as measured by average number of visits per student attending.

5.3.1 Client satisfaction: getting students in the door. The percent of eligible students attending (Fig. 5) is predominantly influenced by marketing. Positive student experiences and instructor, as well as parent, messaging to attend tutoring are forms of marketing that help bring students to the centre and that create a sense of normalcy regarding help-seeking. Beyond marketing, the services offered should meet student needs. The percentage of visiting students also benefits from a convenient physical space location.

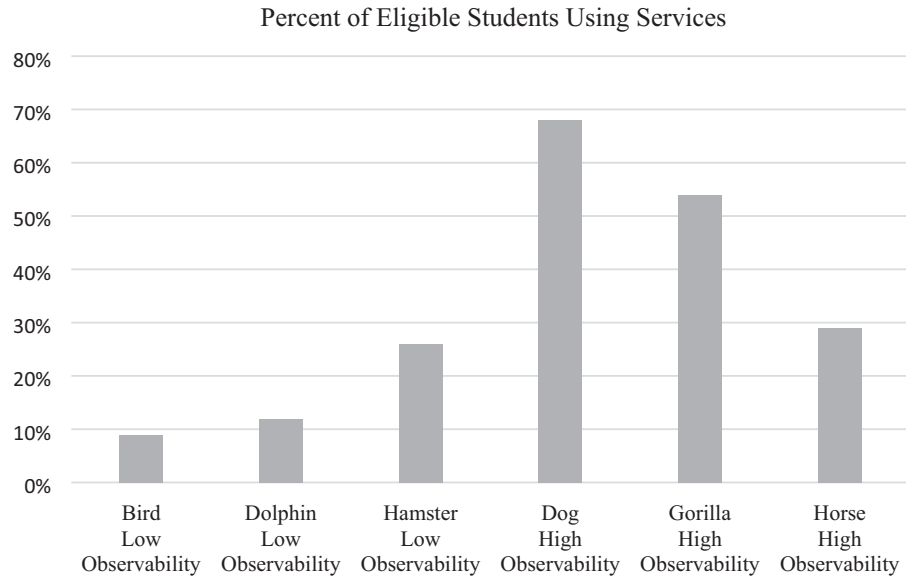

FIG. 5 Client satisfaction measure as percent of eligible students using services for centres sorted by observability.

Compared with other institutions, Hamster's percent of eligible students using their services fell around the middle for the six centres (Fig. 5). The primary marketing tool was instructor announcements, which may be an effective form of marketing since instructors have a captive class audience and are best positioned to recommend tutoring. Hamster had a strong network among instructors as instructor offices were located on the same building floor as the centre tutors where GAs and the centre's two Co-Directors 'work[ed] closely' with instructors and coordinators. This strong relationship served as a reminder of the centre's services and perhaps gave the centre a sense of legitimacy as instructors were familiar with centre leadership and tutors. Furthermore, Hamster's relationship with the campus academic support group aided in the convenience of the physical location and contributed to offering services that met student needs. The mathematics centre was accessible to students, located within a 5-min walk of the middle of campus, and close to mathematics classrooms. The centre also reserved a table to be used for mathematics tutoring appointments offered by the academic support group. Students did not have to go out of their way to attend tutoring, and there was one central location for all mathematics tutoring. Furthermore, Hamster's drop-in services complemented the academic support group appointments. In addition, Hamster's hours of operation were based on student demand, and courses served focused on difficult and higher enrolment service courses.

Bird and Dolphin both had a low percent of eligible students using services compared with other institutions (Fig. 5). In order to pinpoint specific areas for improvement, centres can survey eligible students who are not using their services. While Bird did not conduct such a survey, Dolphin found many students did not attend tutoring because 'they weren't comfortable asking for help.' This indicates the centre had an image issue. Students may have been concerned about being judged by tutors, peers or instructions. To lower the stigma associated with tutoring and improve their image, Dolphin could look to their marketing. Specifically, Dolphin could seek to normalize tutoring with personal testimonies of students who have used the services. Video testimonials of senior students or tutors

who have attended tutoring would demonstrate that even successful students use tutoring, and it is not shameful. In addition, Dolphin may need to improve the instructors' image of the centre. As seen from Hamster, instructors can be positive marketing resources for the centre, but Dolphin reported instructors were sending messaging (either explicit or implicit) that tutoring was only for struggling students. Students may have avoided attending tutoring to prevent being seen or identifying as being 'bad at mathematics'. The centre could partner with instructors to ensure they are sending positive messaging, are making in-class announcements and are referring all, not just struggling, students to tutoring. The centre may find providing instructors with research-based evidence of the benefits of tutoring or student testimonials and local data may improve instructors' image of tutoring. In addition, improving the departmental administration's image of tutoring could help with top-down tutor messaging to instructors.

5.3.2 Client satisfaction: retention. There are also many factors which impact average visits per students attending. The physical space (output) in terms of adequate seating and square foot per student, as well as comfort influence the students' experience. In addition, long wait times due to an insufficient number of tutors would lower satisfaction. The most important factor influencing students' desire to return may be their interactions with the tutors. The behaviours of tutors depend partly on a centre's ability to hire and manage tutors (input). Centres unable to be selective in their hiring or lacking the ability to manage tutors are unable to alter tutors' behaviours. In addition, centres need effective processes for hiring, training and evaluating tutors in order to promote positive student experiences. Surveying students who have attended tutoring helps understand which factors are positively or negatively impacting client satisfaction.

While Bird had difficulties getting students in the door (low percent of eligible students visiting), once students tried tutoring, they kept coming back (high average visits per student attending). Bird had an average of 13.98 visits per student attending, high compared with other institutions (Fig. 6). We examined their centre for factors contributing to their success. Bird's director reported generally positive feedback on student surveys with common praises including '(1) The [centre] provides an active learning experience—students are not given answers but are forced to learn; (2) Tutors are welcoming, motivating and encouraging; and (3) Tutors provide helpful perspectives when learning.' Given the common praises, the high average visits per student is likely due to tutor hiring, training and evaluation.

Bird's director directly hired tutors and reported there were no difficulties finding qualified tutors. This allowed the director to select applicants based on desired characteristics. Criteria used to select tutors included strong communication skills and their own experience learning. Student feedback that tutors are 'welcoming, motivating and encouraging', as well as 'provide helpful perspectives' align with the hiring criteria. In addition, the director hired tutors open to training and served as their manager which put her in a position to train and enforce tutor behaviours aligned with the centre's mission, 'to provide support for active learning'. In addition to other training, Bird's director spent 2–3 h per semester on mathematics active learning pedagogy, such as questioning and collaborative problem solving. Tutors were provided with many opportunities for feedback from both peers and the director. During training sessions, tutors performed simulated tutoring for feedback and were given feedback on end-of-semester self-evaluations. Tutors not meeting expectations for tutor effectiveness would meet with the centre leader to discuss 'how their behaviour could shift, what approaches they should take'. If the tutors did not improve, they were paired with a more expert tutor and worked fewer hours. Achievement of the centre's mission is directly reflected in student feedback that 'the [centre] provides an active learning experience—students are not given answers but are forced to learn'. The training tutors received emphasized active

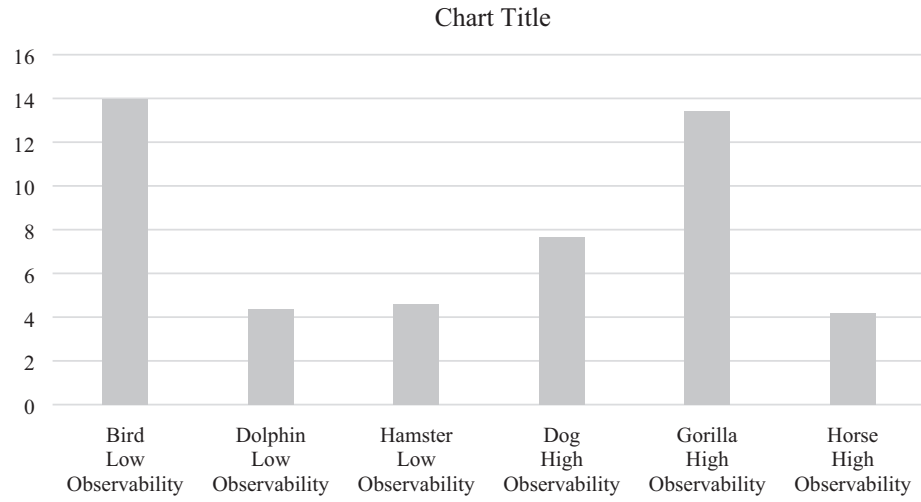

FIG. 6 Client satisfaction measure as average visits per student attending for centres sorted by observability.

learning, mechanisms for tutor feedback regarding their tutor performance contributed to positive client satisfaction.

Hamster and Dolphin had 4.58 and 4.35 visits per student attending, respectively (Fig. 6), which was low compared with other institutions. Hamster and Dolphin each had unique situations regarding student satisfaction surveys. Hamster did not conduct student satisfaction surveys; however, anecdotally, two common complaints were (1) there was not a tutor who could help the student with a certain subject at a given time and (2) 'tutors sit around and don't do anything, only give answers rather than help, and don't seem to care about helping.' These complaints indicate issues with tutor scheduling and communicating hours for specific courses with students. However, perhaps more important, complaints about tutor interactions likely stemmed Hamster's co-directors' inability to directly hire, mange, tutor and evaluate tutors.

Hamster was unable to directly hire or supervise tutors as tutor assignments were based on teaching staff contracts and GA assignments. Tutors were 'typically well qualified, though they are not always familiar with subjects that they have not taken', leading to each tutor tutoring a different subset of courses. The centre's inability to hire tutors meant they could not seek out tutors with desirable characteristics. Had the centre been able to directly hire tutors, they may have been able to more equally distribute the mathematics qualifications of tutors to ensure tutoring was offered for all courses during open hours. In addition, because tutors were not managed by the co-directors, there were no mechanisms in place to evaluate or report tutors not meeting expectations. For example, centre leadership reported, 'ideally a tutor would at least acknowledge a student coming in and ask if they need assistance (once they're seated at settled).' However, the tutors did not meet this ideal 'at all.' Instead, 'tutors sit at tables and work on their own work or play on their phone and occasionally look around to see if anyone has their hand up.' Hamster's co-directors had no means of improving tutor performance. Furthermore, the process for training tutors at Hamster was minimal, including only how the tutoring worked. Tutors were provided online course review material, but its use was not required. The centre's inability to hire, manage and train tutors likely contributed to poor service and low return rates.

Despite Dolphin's low rate of student return, end-of-appointment student surveys completed were typically complementary, with rare complaints about centre hours or a tutor struggling with a mathematics problem. In addition, as this was the centre's first year of operation, tracking the return rate longitudinally could be used to track improvement. To address complaints of tutor knowledge, the centre could examine tutor hiring and training used in the centre. Tutors were hired based on professor recommendation and mathematics honour society membership, likely putting them top among their peers in terms of content knowledge. However, the director could introduce processes to help tutors review content prior to session, creating a procedure by which students inform the tutor of topics they wish to cover in the session, or train tutors in how to either use resources or request help from an instructor, as needed. It may also be beneficial to set student expectations ahead of time that tutors are not instructors and may not have the answers right away but will help students work towards solving the problem together and model using resources. This type of assistance aligns with two of the centre's guiding principles: active learning and to create independent learners. Further, as a centre with low stability, Dolphin can also focus on institutional legitimacy. Ensuring centre services and processes align with the centre mission as key aspect of their performance assessment.

6. Summary and Implications for Practice

By using Lee & Nowell's (2015) framework adapted for mathematics centres, centres can get a much fuller picture of their performance. Previous evaluations for centres have focused primarily on outputs (number of visits), behavioural change outcomes (correlating visits to grades) and client satisfaction outcomes (student surveys). However, Lee & Nowell's (2015) framework takes into account performance dimensions outside of outputs and outcomes. These additional performance dimensions paint a fuller picture of the state of a centre and bring to light the interactions among the dimensions. For example, while centre visits at Dog had a positive effect on grades, observation of what occurred in the centre showed tutors enacted less student-centred pedagogy when the centre was crowded. The centre was using its existing resources efficiently but did not have enough funding for additional tutors. Their relationship with funders in the past had allowed them to request and receive additional funding. Centres can dig beyond outputs and outcomes to understand why these outputs and outcomes come to be.

In addition, previous evaluations have compared tutoring versus no tutoring in order to demonstrate the need for a mathematics centre on campus. Comparing across institutions begins the process of establishing professional standards or ideals to which centres can aspire. For example, Bird can see that compared with other institutions, it has a strong performance on average visits per student attending but low performance on percent of eligible students visiting. These comparisons set benchmarks and allow the centre to turn inward to look for process improvements. Bird can see that students are satisfied once they attend, but the centre may need to improve marketing processes, the centre's image or network with instructors to get students in the door the first time. Centres can use the performance data reported in this study to determine how their centre's performance compares.

Further, centres can use the performance dimension contingencies in Lee & Nowell's (2015) framework to select measures which align with the centre's structure. Centre leaders who are able to spend significant time observing tutors can focus on high observability focal dimensional of the end-to-end process of inputs to outcomes (Fig. 1). These centre leaders are more likely to see how the details of each tutor-student interaction may promote or constrain targeted student behavioural changes. For example, Horse can observe tutor-student interactions to determine if professional standards of interactions (accepted teaching behaviours) are occurring and then adjust training accordingly. While centres with

low observability cannot observe tutor-student interactions, they can use student satisfaction surveys to locate areas for improvement. Hamster's common complaint that tutors ignored students indicates a need for improvement in tutor motivation. Based on their structure, centres can select appropriate performance dimensions from Lee and Nowell's framework which will be useful both for internal improvement and external funding support.

7. Limitations and Future Work

The primary limitations of this work include the cases selected and the data collected. While the cases for this study were selected purposely from the secondary data, the cases in the original study were selected from a convenience sample of math centre leaders who had attended a workshop for mathematics centre leaders, were interested in conducting research and had access to specific student data. These individuals are not representative of the broader mathematics tutor centre community. Glaser & Strauss (1968) argued for selection of 'any groups that will help generate, to the fullest extent, as many properties of the categories as possible, and that will help relate categories to each other and to their properties' (p. 49). Under diverse conditions, different aspects of the theoretical case will be salient. Future work can aim to flush out this work by examining additional cases which differ on key features of those representing in this work (Charmaz, 2000) for improved validity.

Data were also collected for only a single fall semester. The quantitative data did not capture the full annual picture of a centre's performance. In addition, data examining tutor centres' performance over time could be used as an aspect of external validity for data triangulation or as a unit of analysis (Denzin, 1970). As an aspect of external validity, additional sources of data would not only reduce investigator bias but also add additional angles with which to examine the phenomenon. As a unit of analysis, forming a longitudinal or extended case study (Thomas & Myers, 2015) would have allowed for comparing and contrasting or performance within an institution. Future work can collect centre data over several semesters in order to both elaborate on the findings of this study and expand the work to demonstrate who longitudinal within centre data can aid in performance assessment. In addition, future work can examine the unit of analysis to be the broader university, conducting a case study on the collective (Denzin, 1970). Tutor centres do not exist within a vacuum and are subject to institutional and societal influences. One could examine the processes by which the mathematics department culture or university structure impacts a tutor centre. Finally, additional studies are needed which utilize interview and survey techniques to collect data directly from students, tutors, faculty and administration, as well as researcher observations regarding tutor centre operations to add to depth of data and validity.

Funding

National Science Foundation Conference (DUE: 2645086).

Acknowledgements

Collaboration at National Science Foundation Conferences was the inspiration for this paper. The opinions expressed in this paper do not necessarily reflect the position, policy or endorsement of the supporting agency. We would also like to thank Cynthia Farthing and Daniel Moritz for their contributions to early phases of this research program.

References

Akingbola, K. (2012) Context and nonprofit human resource management. *Adm. Soc.*, 45, 974–1004.

Arcavi, A. & Schoenfeld, A. H. (1992) Mathematics tutoring through a constructivist lens: the challenges of sense-making. *J. Math. Behav.*, 11, 321–335.

Ball, D.L., Hill, H.C. & Bass, H. (2005) Knowing mathematics for teaching: who knows mathematics well enough to teach third grade, and how can we decide? *Am. Educ.*, 29 14–17, 20–22, 43–46.

Ball, D. L., Thames, M. H. & Phelps, G. (2008) Content knowledge for teaching: what makes it special? *J. Teach. Educ.*, 59, 389–407.

Berry, E., Mac an Bhaird, C. & O'Shea, A. (2015) Investigating relationships between the usage of mathematics learning support and performance of at risk students. *Teach. Math. Its Appl.*, 34, 194–204.

Booth, D. (2003) Open learning resource center at the University of Huddersfield in "math support for students". LTSN Mathstream. http://mathcentre.ac.uk/resources/mathsteam/booth.pdf (accessed 18 March 2021). 16–17.

Bradshaw, P. (2009) A contingency approach to nonprofit governance. *Nonprofit Manag. Leadersh.*, 20, 61–81.

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qual. Res. Psychol.*, 3, 77–101.

Bressoud, D. M., Mesa, V. & Rasmussen, C. L. (2015) *Insights and Recommendations from the MAA National Study of College Calculus*. MAA Press.

Byerley, C., Moore-Russo, D., James, C., Rickard, B., & Mills, M. (2019) Defining the varied structures of tutoring centers: laying the foundation for future research. *The Twenty-Second Annual Conference on Research in Undergraduate Mathematics Education*. Oklahoma City, OK.

Byerley, C., James, C., Moore-Russo, D., Rickard, B., Mills, M., Heasom, W., Oien, J., Farthing, C., Burks, L., Ferreira, M., Mammo, B., Moritz, D. (2020) Characteristics and evaluation of ten tutoring centers. *The Twenty-Third Annual Conference on Research in Undergraduate Mathematics Education*. Boston, MA.

Byerley, C., Johns, C., Moore-Russo, D., Rickard, B., James, C., Mills, M., Mammo, B., Oien, J., Burks, L., Heasom, W., Ferreira, M., Farthing, C., Moritz, D. (2021) (under review) Towards research-based suggestions for organizational decisions mathematics tutoring centre leader make. (under review).

Byerley, C., Campbell, C. & Rickard, B. (2018) Evaluation of impact of calculus center on student achievement. *Proceedings of the Twenty-first Annual Conference on Research in Undergraduate Mathematics Education*. San Diego, CA, pp. 816–825.

Carmen, J. G. (2011) Understanding evaluation in nonprofit organizations. *Public Perform. Manag. Rev.*, 34, 350–377.

Carroll, C. & Gill, O. (2012) An innovative approach to evaluation the University's mathematics learning centre. *Teach. Math. Its Appl.*, 31, 199–214.

Charmaz, K. (2000) Grounded theory: objectivist and constructivist methods. *The Handbook of Qualitative Research* (N. K. Denzin & Y. Lincoln eds). Thousand Oaks, CA: Sage Publications, Inc., pp. 509–535.

Charmaz, K. (2014) *Constructing Grounded Theory*. London: SAGE Publications.

Chi, M. T. (1996) Constructing self-explanations and scaffolding explanations in tutoring. *Appl. Cogn. Psychol.*, 10, S33–S49.

Child, J. (1972) Organizational structure, environment and performance: the role of strategic choice. *Sociology*, 6, 1–22.

Cooney, T. J., Shealy, B. E. & Arvold, B. (1998) Conceptualizing belief structures of preservice secondary mathematics teachers. *Math. Educ. Res. J.*, 29, 306–333.

Croft, A. C. (2009) Towards a culture of data collection and analysis in mathematics support centres. Queensland, Australia. http://www.sigma-cetl.ac.uk/index.php?section=86 (accessed 18 March 2021).

Croft, A. C., Grove, M. & Bright, D. (2008) A resource and activity centre for mathematics students beyond their transition to higher education. *MSOR Connections*, 8, 11–16.

Denzin, N. K. (1970) *The Research Act in Sociology: A Theoretical Introduction to Sociological Method*. London: Butterworth.

Donaldson, L. (1996) *For Positivist Organization Theory: Proving the Hard Core*. UK: SAGE Publications.

Dowling, D. & Nolan, B. (2006) Measuring the effectiveness of a maths learning support centre—the Dublin City University experience. *CETL-MSOR Conference 2006 Conference Proceedings Published by The Maths, Stats, and OR Network Printed by Central Printing Services*. University of Birmingham, pp. 51–54. https://

www.sigma-network.ac.uk/wp-content/uploads/2016/11/CETL-MSOR-Proceedings2006.pdf (accessed 18 March 2021).

Dzator, M. & Dzator, J. (2020) The impact of mathematics and statistics support at academic learning centre, Central Queensland University. *Teach. Math. Its Appl.*, 39, 13–28.

Eisenhardt, K. M. (1989) Building theories from case study research. *Acad. Manag. Rev.*, 14, 532–550.

Gill, O. & O'Donoghue, J. (2007) Justifying the existence of mathematics learning support: measuring the effectiveness of a mathematics learning centre. *Paper Presented at Adults Learning Mathematics 14 (ALM 14)*, Limerick, June 26–29, pp. 154–164.

Gillard, J., Robathan, K. & Wilson, R. (2011) Measuring the effectiveness of a mathematics support service: an email survey. *Teach. Math. Its Appl.*, 30, 43–52.

Glaser, A. L. & Strauss, B. G. (1968) *Nursing Research*. Chicago: Aldine Publications.

Graesser, A. C., Person, N. K. & Magliano, J. P. (1995) Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Appl. Cogn. Psychol.*, 9, 495–522.

Herman, R. D. & Renz, D. O. (2008) Advancing nonprofit organizational effectiveness research and theory. *Nonprofit Manag. Leadersh.*, 18, 399–415.

Jacob, M. & Ni Fhloinn, E. (2019) A quantitative, longitudinal analysis of the impact of mathematics support in an Irish university. *Teach. Math. Its Appl.*, 38, 216–229.

Karsenty, R. & Sherin, M. G. (2017) Video as a catalyst for mathematics teachers' professional growth. *J. Math. Educ.*, 20, 409–413.

King, N. & Brooks, J. (2018) Thematic analysis in organisational research. *The SAGE Handbook of Qualitative Business and Management Research Methods: Methods and Challenges* (C. Cassell, A. L. Cunliffe & G. Grandy eds). UK: SAGE, pp. 219–236.

Kyle, J. (2010) Affordability, adaptability, approachability and sustainability. *Responding to the Mathematics Problem* (C. M. Marr & M. J. Grove eds). St Andrews: Maths, Stats & OR Network, pp. 103–104.

Lawson, D. A. (2015) Mathematics support at the transition to university. *Transitions in Undergraduate Mathematics Education* (M. J. Grove, A. C. Croft, J. Kyle & D. A. Lawson eds). Birmingham: The University of Birmingham & Higher Education Academy, pp. 39–56.

Lawson, D., Grove, M. & Croft, T. (2020) The evolution of mathematics support: a literature review. *Int. J. Math. Educ. Sci. Technol.*, 51, 1224–1254.

Lee, C. & Nowell, B. (2015) A framework for assessing the performance of nonprofit organizations. *Am. J. Eval.*, 36, 299–319.

Lepper, M. R. & Woolverton, M. (2002) The wisdom of practice: lessons learned from the study of highly effective tutors. *Improving Academic Achievementt: Impact of Psychological Factors on Education* (J. Aronson ed.). San Diego, CA, USA: Academic Press, pp. 135–158. https://doi.org/10.1016/B978-012064455-1/50010-5.

Mac an Bhaird, C., Gill, O., Jennings, K., Ni Fhloinn, E., O'Sullivan, C. (2011) The Irish mathematics support network: its origins and progression. *AISHE-J*, 3, 51.1–51.14.

MacGillivray, H. L. & Croft, A. C. (2011) Understanding evaluation of learning support in mathematics and statistics. *Int. J. Math. Educ. Sci. Technol.*, 42, 189–212.

Matthews, J., Croft, A. C., Lawson, D. A. & Waller, D. (2013) Evaluation of mathematics support centres: a literature review. *Teach. Math. Its Appl.*, 32, 173–190.

Mills, M., Rickard, B. & Guest, B. (2020) Survey of mathematics tutoring centers in the USA. *Int. J. Math. Sci. Educ. Technol.*, 1–21. Published online. https://doi.org/10.1080/0020739X.2020.1798525.

Mintzberg, H. (1980) Structure in 5's: a synthesis of the research on organization design. *Manag. Sci.*, 26, 322–341.

Mitchell, J. C. (1983) Case and situation analysis. *Sociol. Rev.*, 31, 187–211.

Ni Fhloinn, E. (2009) The role of student feedback in evaluating mathematics support centers. *Proceedings of the CETL-MSOR Conference 2009*. Milton Keynes, UK, pp. 94–98 http://doras.dcu.ie/21665/1/NiFhloinn2009.pdf (accessed 18 March 2021).

O'Sullivan, C., Macan Bhaird, C., Fitzmaurice, O. & Ni Fhlionn, E. (2014) *An Irish Mathematics Learning Support Network (IMLSN) Report on Student Evaluation of Mathematics Learning Support: Insights from a*

Large Scale Multi-Institutional Survey. Ireland: National Centre for Excellence in Mathematics and Science Teaching and Learning.

Parsons, S., Croft, A. C. & Harrison, M. (2011) Engineering students' self-confidence in mathematics mapped onto Bandura's self-efficacy. *J. Eng. Educ.*, 6, 52–61.

Rickard, B. & Mills, M. (2018) The effect of attending tutoring on course grades in Calculus I. *Int. J. Math. Educ. Sci. Technol.*, 49, 341–354.

Rojas, R. R. (2000) A review of models for measuring organizational effectiveness among for-profit and nonprofit organizations. *Nonprofit Manag. Leadersh.*, 1, 97–104.

Rossi, P. H., Lipsey, M. W. & Freeman, H. E. (2004) *Evaluation: A Systemic Approach*, 7th edn. London: SAGE Publishing.

Rylands, I. J. & Shearman, D. (2018) Mathematics learning support and engagement in first year engineering. *Int. J. Math. Educ. Sci. Technol.*, 49, 1133–1147.

Santagata, R. & Guarino, J. (2011) Using video to teach future teachers to learn from teaching. *ZDM Math. Educ.*, 43, 133–145.

Santagata, R. & Yeh, C. (2014) Learning to teach mathematics and to analyze teaching effectiveness: evidence from a video and practice-based approach. *J. Math. Teach. Educ.*, 17, 491–514.

Schoenfeld, A. H. (2017) Use of video in understanding and improving mathematical thinking and teaching. *J. Math. Teach. Educ.*, 20, 415–432.

Schon, D. A. (1987) *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. San Fransisco, CA: Jossey-Bass.

Smith, E. (2008) Pitfalls and promises: the use of secondary data analysis in educational research. *Br. J. Educ. Stud.*, 56, 323–339.

Staddon, L. & Newman, L. (2003) The education drop-in centre at the University of Glamorgan, in "maths support for students". LTSN mathstream, 22–23.

Thomas, G. & Myers, K. (2015) *The Anatomy of the Case Study*. Los Angeles: SAGE Publications.

Thompson, A. G. (1992) Teachers' beliefs and conceptions: a synthesis of the research. *Handbook of Research on Mathematics Teaching and Learning* (D. A. Grouws ed). New York: Macmillan, pp. 127–146.

Toepler, S. & Anheier, H. K. (2004) Organizational theory and nonprofit management: an overview. *Future of Civil Society* (A. Zimmer & E. Prillar eds). Wiesbaden: VS, pp. 253–270.

Van Es, E. A. & Sherin, M. G. (2008) Mathematics teachers' "learning to notice" in the context of a video club. *Teach. Teach. Educ.*, 24, 244–276.

Van Es, E. & Sherin, M. G. (2010) The influence of video clubs on teachers' thinking and practice. *J. Math. Teach. Educ.*, 13, 155–176.

Wadongo, B. & Abdel-Kader, M. (2014) Contingency theory, performance management and organisational effectiveness in the third sector: a theoretical framework. *Int. J. Product. Perform.*, 63, 680–703.

Wilkins, L. (2015) Maybe we could just count the boxes of chocolates? Measuring the impact of learning development mathematics support for undergraduate students. *JALL*, 9, A91–A115.

Xu, Y., Hartman, S., Uribe, G. & Mencke, R. (2001) The effects of peer tutoring on undergradaute examination scores in mathematics. *J Coll Read*, 32, 22–31.

Yin, R. K. (2018) *Case Study Research and Applications: Design and Methods*, 6th edn. Los Angeles, CA: SAGE Publications, Inc.

Carolyn Johns is the Assistant Director of the Mathematics and Statistics Learning Center at Ohio State. Her research interests broadly include mathematics tutor practices, tutor training and teacher education. Prior to working at The Ohio State University, she received an undergraduate degree from Earlham College and received a Ph.D. from The Ohio State University in August 2019.

Cameron Byerley is an assistant professor at the University of Georgia. Formally, she co-directed a university mathematics tutoring center. Her research interests include modelling students' and citizens' mathematical thinking.

Deborah Moore-Russo, Ph.D., is the First-Year Mathematics Director at the University of Oklahoma. Her primary research interests include the ways that mathematics topics are conceptualized, represented, visualized and communicated. She is also interested in how mathematics departments support first-year mathematics courses through tutoring centers, course coordination and other efforts.

Brian Rickard earned a BS in Mathematics, MEd in Higher Education and Ph.D. in Educational Statistics and Research Methods at the University of Arkansas. He is a teaching assistant professor and coordinator of Finite Mathematics in the Department of Mathematical Sciences at the University of Arkansas. He conducts research in undergraduate mathematics education and university tutoring centers.

Janet Oien is the Co-Director of the Calculus Center at Colorado State University. Her research interests broadly include mathematics tutor practices, tutor training and preparation of secondary mathematics teachers. She received her undergraduate degree and M.S. degree in Mathematics with an outside concentration in Education from Colorado State University in 2006.

Linda Burks is the Director of the Mathematics Learning Center and a math lecturer at Santa Clara University. She received her undergraduate degree in mathematics from Virginia Tech; she earned her M.S. in mathematics and Ph.D. in mathematics education from the University of Michigan. Her research interests include mathematics tutor practices and cognitive self-regulation.

Carolyn James, Ph.D., is the Calculus Coordinator and Mathematics Resource Center Supervisor at University of Portland. She received her undergraduate degree in mathematics at Carleton College and her Ph.D. in mathematics education from Portland State University. Her research interests include barriers and drivers of instructional change, the relationship between instructor beliefs and practice and undergraduate mathematics tutoring.

Melissa Mills is a teaching associate professor and the Director of the Mathematics Learning Success Center at Oklahoma State University. Her research interests broadly include mathematics tutor practices and the teaching and learning of mathematical proof at the undergraduate level.

William Heasom, MA, MSCE, is assistant to the Director of the Mathematics Learning and Resource Center at Villanova University He also serves as collaborator with the IARPA FOCUS (forecasting counterfactuals

in uncontrolled settings) project at University of Pennsylvania where he helped develop training materials for intelligence analysts. His research interests concern applying mathematical and agent-based modelling techniques to the analysis and design of hydrologic, ecological and learning systems.

Melissa Ferreira is the Director of the Mathematics Learning and Resource Center at Villanova University. Her research interest is in university mathematics tutoring. She received undergraduate and masters' degrees in Mathematics from Villanova University.

Behailu Mammo is a professor of mathematics and Director of the Mathematics Tutorial Center at Hofstra University, NY. His research interest focuses on number theory and examining the impact of tutoring in improving pre-service teachers' dispositions and self-efficacy. In May 2005, he received his Ph.D. from Temple University, PA. His bachelor's and master's degrees are from Ethiopia, East Africa.