

Ghost hunting with lollies, chess and Lego: appreciating the ‘messy’ complexity (and costs) of doing difficult research in education

Linda J. Graham · Linda Buckley

Received: 8 December 2012 / Accepted: 29 November 2013
© The Australian Association for Research in Education, Inc. 2013

Abstract This paper contributes to conversations about the funding and quality of education research. The paper proceeds in two parts. Part I sets the context by presenting an historical analysis of funding allocations made to Education research through the ARC’s Discovery projects scheme between the years 2002 and 2014, and compares these trends to allocations made to another field within the Social, Behavioural and Economic Sciences assessment panel: Psychology and Cognitive Science. Part II highlights the consequences of underfunding education research by presenting evidence from an Australian Research Council Discovery project that is tracking the experiences of disaffected students who are referred to behaviour schools. The re-scoping decisions that became necessary and the incidental costs that accrue from complications that occur in the field are illustrated and discussed through vignettes of research with “ghosts” who don’t like school but who *do* like lollies, chess and Lego.

Keywords Research funding and quality assessment · Qualitative and mixed-methods research · ERA · Australian Research Council

Part I: the funding of research in Australia

In recent years, higher education and research has been viewed as a key driver in the development of an innovative “knowledge-industry” aimed at ensuring the competitive advantage of individual nation-states in a rapidly changing world

L. J. Graham (✉)
Faculty of Education, Queensland University of Technology, Rm 326, A Block, Victoria Park Road,
QLD 4059, Australia
e-mail: l2.graham@qut.edu.au

L. Buckley
Department of Linguistics, Macquarie University, North Ryde, NSW 2109, Australia

order (Peters and Besley 2006). Australia is no exception to this trend. While the higher education sector has been witness to significant change since the 1960s, externally-set imperatives have increased in the last decade with successive federal governments deploying targeted “incentive funding” to direct the focus, style and quality of university teaching and research (Marginson 2007). One example of such incentive-based funding is the Australian Competitive Grants (ACG) scheme distributed principally via the Australian Research Council (ARC) and the National Health and Medical Research Council (NHMRC). The ARC funds research across all disciplines with the exception of clinical medical research and dentistry,¹ which are the domain of the NHMRC. Academics in disciplines focusing on health-related issues, such as science, psychology and linguistics can apply to both research councils; however, research in education is almost exclusively funded by the ARC.

Prior to 1989, the allocation of public research funding was managed via a university-based system funded by the Australian government through regular operating grants (Bazeley 2003). During the early 1990s, in what has become known as “the Dawkins era”, this funding was re-appropriated in order to establish a centralised system, which was “designed to ensure that research was funded only on the basis of excellence with equal rigour across and within universities” (Bazeley 2003, p. 258). The ACG system effectively created a competition for research funding which, together with the expectation that all academics would engage in research, produced a situation where both the number of proposals and the quality of submissions soon exceeded the limited funding available (Bazeley 2003).

Two decades on and not only does ACG funding now represent a major source of university prestige and research income, but its procurement has significant implications for academic career success. Since the introduction of the Excellence for Research in Australia (ERA) research quality assessment exercise in 2010, the ACG process has become increasingly competitive and even more high-stakes. The inclusion of all university academic staff in the ERA count has increased pressure on all academics to not only become research-active but to apply for and, if possible, secure ACG funding. In some universities, submitting an ACG application within the first 3 years of employment is a condition of successfully passing probation and/or achieving promotion. Not surprisingly, while there have been significant increases in funding allocations to the ARC and NHMRC, each funding body has struggled to maintain success rates against the rise in research funding applications (Allen Consulting Group 2008).

Discovering the impacts on education research

The ARC has a number of competitive grant schemes, the largest and most prestigious of which is the Discovery Projects scheme, where the greatest assessment weighting (40 %) focuses on researcher track record. Increased pressure on academics to secure research funding saw application rates rise by more than a third in the 7 years between 2002 (Discovery’s year of inception) and 2011 (the

¹ This distinction is not always upheld. Between 2009 and 2013, ARC Future Fellowships were open to clinical medical researchers and dentistry.

Table 1 ARC Discovery application and success rates 2002–2014

Year	No. of proposals	Success rate (%)
2014	3,534	19.9
2013	3,425	21.4
2012	3,544	21.95
2011	4,230	22.0
2010	4,068	22.7
2009	4,152	20.4
2008	4,112	21.4
2007	4,033	20.4
2006	3,048	24.5
2005	3,441	30.9
2004	3,260	27.0
2003	3,601	25.8
2002	3,090	23.2

year with the highest number of proposals on record). In recent years, the ARC has sought to ration available funding by cutting allocations to successful projects. For example, although there was a 34.55 % increase in funding awarded through Discovery in the 2002–2014 period, the real cost of research also grew, leading to a decrease in the average amount awarded per grant from 75 % of funding requested in 2002 to 63.8 % in 2014.² Whilst application rates dropped back to 2003 levels from 2012, perhaps reflecting increased university scrutiny and the internal culling of proposals that are very unlikely to be successful, success rates have continued to decline (see Table 1).

Overall success rates tend to mask cross-discipline variability. The ARC groups research disciplines under five panels: Biological Sciences and Biotechnology (BSB); Engineering, Mathematics and Informatics (EMI); Humanities and Creative Arts (HCA); Physics, Chemistry and Earth Sciences (PCE); and Social, Behavioural and Economic Sciences (SBE). While projects fitting the description of education research can be and are submitted to other panels, such as HCA (2014 success rate 21.6 %), the majority go to SBE which achieved the lowest success rate of 19.9 % in 2014. Given that SBE is itself a composite of research disciplines, it is worth asking whether the lower success rate has impacted more heavily on the field of education relative to the other fields of research within the SBE panel like, for example, Psychology and Cognitive Science.³

Statistics on success rates and amounts requested are not published for specific fields of research, however, a comparison of the number and value of projects awarded under Education in 2002 with the number and value of projects awarded

² Note that this did drop to an average of 55.3 % in 2011 when funding for the current study was awarded. While there has been an increase in individual project amounts awarded between 2011 and 2014, and thus a rise in the percentage of funding allocated per project relative to funding requested, the effect of this increase has been to reduce the number of successful proposals in Education overall (from 19 to 12).

³ Education and Psychology have been paired in previous considerations of the ARC funding process (see Marsh et al. 2008).

under Education in 2014 reveals that there has been a decline in the proportion of research funding allocated.⁴ In 2002, there were 25 successful Discovery proposals in Education totalling \$3,119,500 or 1.63 % of the total funding pool. By 2014, however, there was *less than half* the number of successful proposals (12 in total). These amounted to \$3,428,699 or 1.33 % of the total funding pool, resulting in an 18.4 % decrease in the share of total ARC Discovery funding netted by Education. While the SBE success rate of 19.9 % was again one of the lowest in 2014, the downward trend experienced by Education was not shared by our peers in Psychology and Cognitive Science (see Fig. 1 below). In 2002, there were 35 successful proposals in Psychology and Cognitive Science totalling \$6, 378,258; a net dollar amount that represented 3.33 % of the total funding pool and which was more than *double* that awarded to Education. By 2014, however, this had increased to 39 successful proposals netting \$14,033,809 or 5.45 % of total ARC funding for Discovery Projects. This amount was more than *four times* that awarded to Division 13 Education and one that represented a 63.66 % increase in the share of total ARC Discovery funding received by Division 17 Psychology and Cognitive Science.

Indeed, ARC Discovery funding to Psychology and Cognitive Science more than *doubled* in the 2002–2014 period with an increase of more than \$7 million; whereas Education received only \$309,199 more in 2014 than it did in 2002 (see Table 2 below). Recall that the real cost of research also grew during this period. Furthermore, projects in Education were awarded an average amount of \$202,391 in 2014, compared to \$359,841 for Psychology and Cognitive Science (see Table 3). Given the complexities and cost of conducting research in schools discussed later in this paper, these differences have a serious dampening effect on research relating to a very large and important sector of our economy. It is also important to bear in mind that Education research is almost exclusively funded by the ARC and that Psychology and Cognitive Science also secures a significant share of funding from

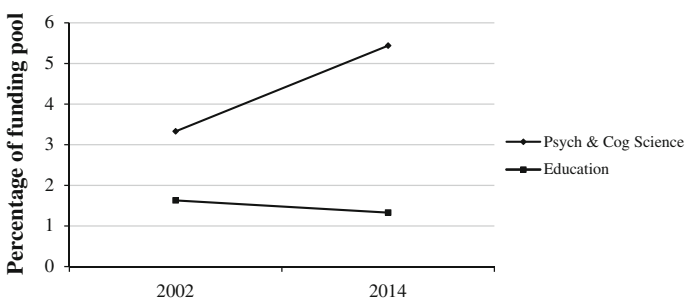


Fig. 1 Percentage of ARC Discovery funding awarded to Psychology and Cognitive Science relative to Education between 2002 and 2014

⁴ During the period under study the ARC changed from Research Fields, Courses and Disciplines (RFCD) codes to Divisions and Field of Research (FoR) codes. This analysis has therefore compared projects awarded under the parent RFCD Code for Education (330000) and Psychology (380000) in 2002 with Division 13 Education and Division 17 Psychology and Cognitive Science, in 2014.

Table 2 Real and percentage change in funding for Discovery, comparing divisions 13 and 17

Division	2002	2014	Percentage change (%)
Total Discovery funding pool	\$191,473,765	\$257,632,541	34.55
13 Education	\$3,119,500	\$3,428,699	9.91
17 Psychology and Cognitive Science	\$6,378,258	\$14,033,809	120.03

Table 3 Average project funding awarded for divisions 13 and 17

Division	2002	2014	Percentage change %
13 Education	\$124,780	\$202,391	62.2
17 Psychology and Cognitive Science	\$182,235	\$359,841	97.46

the NHMRC. The funding gap between these important disciplines is therefore much larger than indicated here.

Received wisdom positions these discrepancies as the result of merit and the superiority of ‘scientific’ method, which feeds a general assumption that academics in the cognitive and behavioural sciences have received better research training and are thus more able to write proposals of a competitive standard. Having served as an ARC Assessor for some years, predominantly in the area of Special Education and Disability, I have had the opportunity to assess proposals from both fields and would argue that there is an equal degree of good and bad in each. As opposed to being an indicator of quality or a lack thereof, the disproportionate success of Psychology and Cognitive Science proposals, relative to those from Education, may instead reflect the disproportionate under-representation of Education researchers on the SBE panel of experts and the dearth of qualitative researchers in the ARC College of Experts more generally. It may simply be that Education research, itself a diverse and complex area of study involving multiple sub-disciplines and methods, is not well understood by the researchers who are making funding decisions on behalf of *all* the disciplines within the Social, Behavioural and Economic sciences.

There are a number of implications that flow from both the shortage in funding and its concentration, the most obvious being that Australia risks strangling the development of future researchers who have the ability to conduct high quality research in an important, yet complex and poorly understood field. Another equally important consideration is that quality research in education does not come cheaply and we risk getting what we pay for. The consequence of this is profound both for the country and the field. Although education research is often dismissed as “low-rent” by some university research offices more intent on the efforts of those who may more reliably bring in the money, research in education is of critical importance to our nation’s future. Underfunding research in a field that is already perceived to be lacking in depth will do nothing to improve the quality of work conducted. Indeed, underfunding will make it exceptionally hard for even the most experienced and talented researchers to produce quality results. Further, with some Discovery projects still receiving significantly less than the funding requested, the

ARC's rationing strategy has further exacerbated the problem of grants not covering the full costs of conducting research (Healy 2009). The impact on research in the field of education, which "usually gets the crumbs, irrespective of the quality of the work we do" (Goodyear 2008, p. 4), is high because of the labour intensiveness of education research and the "messiness" of doing research in schools.

This paper aims to contribute to conversations about the inadequacy of funding for and value of research in education by presenting evidence from an ARC Discovery project that is tracking the experiences of students referred to behaviour schools in New South Wales. The following sections outline the re-scoping decisions that became necessary and the challenges involved when attempting to recruit a large enough sample to retain quantitative integrity from the hardest-to-reach parents and young people with just over half of the funds needed to do so. The incidental costs that accrued from complications that occur in the field—which could not be budgeted for, but would more easily have been absorbed under a full cost model of research funding—will be illustrated and discussed through vignettes of research with ghosts who don't like school but who *do* like lollies, chess and Lego.

Part II: the study

New South Wales (NSW) is Australia's largest state comprising one-third of the national population. In 2011, there were 1,128,317 school-aged students, 34 % of whom attended non-government schools (Schools Australia 4221.0, 2011). The government school sector educates 66 % of the K-12 schooling population with over 2,200 schools and 744,392 students (Schools Australia 4221.0, 2011). The NSW government school sector also has the largest parallel special education system in Australia with over 2000 support classes and 113 special schools. These schools, termed "Schools for Specific Purposes" (SSPs), form the most restrictive placement option. Over the last two decades there has been significant increase in their use, mainly due to growth in enrolments under the categories of emotional disturbance and behaviour disorder (Graham and Sweller 2011). Recent research has found that this increase coincided with the construction of a series of "behaviour" schools (Graham et al. 2010), which now account for more than one-third of all special schools in the NSW government schooling sector (Graham 2012). Of significant concern however is the lack of quality research to indicate what effect such responses have on the children involved or what contribution is made to their future educational success. Despite this gap in the research knowledge, the use of special schools and classes for students who engage in disruptive behaviour is growing, while the average age of the students who attend is decreasing (Graham et al. 2010).

The aims of this ARC Discovery project (Graham et al. 2011) therefore were to:

- (1) document how such interventions take form and the ways in which these are perceived by students and school personnel;
- (2) trace student memories of their prior schooling experience and what connection, if any, these students make between these experiences and where they are now;

- (3) track changes in student attitudes, beliefs and behaviour during and after their enrolment in special schools; and
- (4) observe and analyse student re-integration to regular schooling to understand what events lead to positive and negative experiences.

To engage with these aims, the proposed project utilised a cross-sectional mixed-method longitudinal research design (Creswell 2003; Takkashori and Teddlie 1998) with 84 participants aged between 9 and 14 years. The research participants were to be recruited in three groups of 28 from both mainstream and separate special educational settings; including 28 students currently enrolled in separate behaviour schools (the longitudinal group), 28 students *with* a history of severely disruptive behaviour enrolled in mainstream schools, and 28 students enrolled in mainstream schools with *no* history of disruptive behaviour.

The project was scheduled to progress in three phases. In Phase 1, each student was to participate in a series of assessments including the Peabody Picture Vocabulary Test (PPVT-IV, Dunn and Dunn 2007) and the Expressive Vocabulary Test (EVT-2; Williams 2007), plus a semi-structured interview that was designed to tap into students' attitudes to schooling; their perspectives on school work; their views on teachers and teaching; their self-perception and peer-relationships; their experiences of change and dislocation; their future aspirations; and their memories of current and past schooling experiences. Phases 2 and 3 were scheduled at 6 monthly intervals to longitudinally follow-up with the behaviour school group in order to gauge whether their perceptions or experiences changed over the ensuing 12 months.

To help situate individual student data within a broader ecological context, the project design also drew on case-study methodology (Yin 2003). The aim was to recruit the longitudinal group (28 × behaviour school students) from four case study behaviour schools (7 per students per school) which, given the geographic spread between the most appropriate research sites,⁵ would help minimise cost by limiting travel and time in the field. The Research Assistant was to spend 3 weeks immersed in each of these schools, using the time to conduct the second round of interviews with the 28 behaviour school students, as well observe student behaviour and teaching practice in classrooms and playgrounds. In Phase 3, each student in the behaviour school group was to be observed twice during their re-integration in the mainstream environment and participate in a final follow-up interview. In all, the project aimed to conduct 140 student interviews of approximately 30 min duration, 84 language assessments, plus approximately 180 h of general observation in the behaviour schools and 168 h of observation tailing reintegrating students who had returned to mainstream. The proposed, and somewhat conservative, budget for the research was \$197,000; a major component of which was transcription, field-related travel expenses and the planned appointment of an experienced part-time research assistant for the first 2.5 years. The project, one of only 19 successful Education

⁵ Given that there are less than 30 behaviour schools in the state, only limited information can be provided about the participating schools. Each was selected on the basis of size, location and willingness to participate. The maximum class size in a behaviour school is 7 students and most behavior schools have only 3–4 classes (or 21–28 students).

proposals in the 2011 Discovery round, was awarded \$110,000 (55 % of the funding requested); a sum which had immediate implications.

Re-scoping “grand designs”

The first impetus when faced with a situation like this is to reduce scale. Researchers are accustomed to making minor adjustments to the proposed number of interviews or case-study schools without it completely destroying their plans for analysis. In this project however, a 45 % reduction in the number of participants—e.g., from 84 to 46—would have seriously weakened the quantitative half of the research design. To achieve the statistical power necessary to conduct analyses beyond descriptive statistics, the minimum number for each group was 20. However, even reducing the total number of participants from 84 to 60 was not feasible for a number of reasons. Firstly, the project was focusing on an extremely hard-to-reach population and oversampling in the first (and largest) round would be necessary due to the high likelihood of attrition in the subsequent rounds. Secondly, two of the project’s CIs were quantitative researchers with specialist expertise that was pertinent to the aims of the project and which could not be realised through a purely descriptive, qualitative study. Thirdly, a reduction of only 24 participants would fail to achieve the full cost savings required.

Adding to this challenge was a Department of Education and Communities (DEC) requirement that the project recruit only from “pure” behaviour schools. The suggestion was made because “ED/BD” schools (administered by Disability Programs Directorate) require a confirmation of diagnosis in a disability category eligible for support in NSW government schools for entry, whereas behaviour schools (administered by Student Welfare Directorate) require no confirmation of disability (Graham 2012). The then DEC Director of Student Welfare was of the view that this risked confounding the research results because students in “ED/BD” special schools would have mental health issues but students in behaviour schools were simply “disruptive” and may not have participated in any formal assessment or diagnostic process prior to enrolment.⁶ This had no impact on the planned case-study schools—as these *were* “pure” behaviour schools—but it did mean that we would need to include measures that would provide a way of determining whether and how students differed in terms of behaviour and anxiety levels across our three settings. After some deliberation, the measures selected were the State/Trait Inventory Test (Spielberger 1973) and all three versions of the Child Behaviour Check List: Teacher, Parent and Youth Reports (Achenbach 1991). The principal reason for the inclusion of all three CBCL report forms was to examine whether there were any significant between-group differences in how students perceive their own behaviour and whether there were differences when students’ perceptions were compared to those of their parents/caregivers and teachers. The inclusion of the

⁶ Upon interviewing the principals of our case-study behaviour schools, the first author discovered that this was not the case and that students in behaviour schools may have multiple diagnoses, including emotional disturbance, anxiety and autism spectrum disorder.

additional measures was another unanticipated cost but one that was made necessary by structural realities within the NSW government school system.

Given that the principal focus of the project was to track the perspectives *and* experiences of students enrolled in behaviour schools, the decision was made to invest more significantly in the interviews and assessments and to section out the case-study research for a university-sponsored doctoral project (Granite and Graham 2012). The first author also increased her own time contribution (from 0.2 to 0.6 for the first 18 months of the project), as well as injecting \$20 K additional funding from other sources. This not only increased the quality and representativeness of the research but it also enabled the team to absorb additional costs without which the project could not have been realised.

For example, as the project progressed to the recruitment stage it became very clear that achieving even the minimum number of 20 in each group was going to be a significant challenge. Over 500 parent information and consent packs (which included the Achenbach Child Behaviour Check List Parent Report Form and reply-paid envelopes) were distributed to 30 participating schools (7 behaviour schools and 23 mainstream feeder schools), however response rates were low and slow.

Our behaviour school principals explained that they often had difficulty in getting their students to return notes and suggested that we offer a can of soft drink or scratchie tickets as they often did. Obtaining responses from mainstream feeder schools was even more difficult. There were a number of reasons for this, ranging from general research fatigue on the part of schools to hostility towards research involving students with behavioural issues. After some 8 weeks (with only four return consent forms in total), Human Research Ethics approval was sought to offer a double movie voucher (to the value of \$25) to compensate parents for the time required to complete the Child Behaviour Check List Parent Report Form (Achenbach 1991) and return it with their consent form. Some of our schools also offered to call parents to check whether they had received the forms from their child and/or whether they would like the school to post them to their home address. Return of the consent forms picked up as a result and we were successful in recruiting 96 students in total (see Table 2), however, the process of data collection represented another series of hurdles.

Chasing ghosts

While we finally did begin to receive parent consent, it soon became apparent that absenteeism was rife in some schools and that school knowledge of student whereabouts was not very reliable. Data collection became extremely complicated because most schools wanted at least 1 day's notice before we arrived to work with a student, however, because some were very hard to catch, this requirement left us vulnerable to arriving for appointments that had been confirmed the previous day but with no students to work with. Time and again, we were assured by school personnel that the participating students were attending school and would be available at the scheduled time, only to find that they hadn't turned up that day, or that they'd arrived and left before we got there or that they'd had a "meltdown" and were being picked up. We began tracking attendance and discovered that some of

our “ghosts” had not been at school for months. Office staff finally confirmed this. In one instance, the receptionist who answered one of our regular phone queries called over her shoulder to other office staff to ask, “When do we get rid of No. 32? August 31st is his 17th birthday, yeah?” Her advice was that they hadn’t seen No. 32 in months but that he would cease to be their responsibility upon reaching the school leaving age (Table 4).

We never did meet No. 32.⁷ Nor did we meet with more than 10 others for whom we had received parent consent but whom, for a number of reasons, were not available to interview for weeks at a time. While this created logistical difficulties, it was also deeply concerning that the schools often had no idea whether these students were at school and, if not, where they actually were. We began to document this through field-notes (see Table 2 for an explanation of school codes):

Today was a good day for collecting data on the difficulties of doing this research. As you know, I went to School 3B [Western Sydney] at 9 am but was unable to see No. 4. School 12M is across the road so I rang them to ask if No. 65 was at school, after I had expected to see him yesterday but discovered in the afternoon that he wasn’t there... Jane apologised for saying that No. 65 was at school yesterday when he wasn’t. I thanked her and said that I understood, and that this happens sometimes. (It sure pays to keep onside with the office staff!) After a hunt for No. 65 ... they discovered that he wasn’t at school today either. He is apparently a good school attender and it was unusual that he was away for 2 days in a row. They thought he might be sick, but didn’t know for sure. Next I rang School 16M but [the principal] was on the phone. The office lady said that No. 70 was present, unsure about No. 68, and No. 69 was absent. I asked if I could start heading to [outer South Western Sydney] to see them. She said it was fine, but she would need to check with whoever was running the program—she thought it might be Miss Brown. Later, when I was near Liverpool, the office lady rang back to say that Kate is running the program and it’s not convenient to see the students today. Also, that No. 68 is unsettled today. She would like at least a day’s notice before I visit. Next I rang School 1B as we are chasing No. 62 and No. 32. No. 62 had been at school for 10 min this morning, then left. They haven’t seen No. 32 for weeks. Sheila said to hold on and came back to say that we should cross them off our list because they are never going to show up. Next, I rang School 9M [South Western Sydney] and No. 67 was at school AND I was welcome to visit... I drove about 150 km today!
(Linda B, Field Notes, 06.03.12)

Our constant follow-ups meant that we finally did manage to see No. 62 (in May 2012), and we also learned why No. 65 was missing.

⁷ As we had 96 students participating in this project, assigning individually unique pseudonyms is problematic. Whilst it can be argued that assigning numbers rather than names is dehumanizing, for the purposes of this article I am using the code that was assigned to each new consent form. These numbers held special significance for us as they helped to buoy our spirits during an arduous recruitment process, as well as ensure anonymity for our participants.

Table 4 Characteristics and number of participants for each participating school

School ID	School type	No. of participants (% of total)
1B	Behaviour	5 (5.21)
2B	Behaviour	5 (5.21)
3B	Behaviour	1 (1.04)
4B	Behaviour	6 (6.25)
5M	Primary	3 (3.13)
6M	Primary	12 (12.50)
7B	Behaviour	9 (9.38)
8M	Secondary	4 (4.17)
9M	Secondary	3 (3.13)
10M	Secondary	2 (2.08)
11M	Secondary	4 (4.17)
12M	Primary	5 (5.21)
13B	Behaviour	1 (1.04)
14M	Primary	1 (1.04)
15B	Behaviour	6 (6.25)
16M	Secondary	14 (14.58)
17M	Secondary	5 (5.21)
18M	Secondary	9 (9.38)
19M	Primary	1 (1.04)
Total		96 (100)

I have been ringing School 12M this morning to see if No. 65 was at school. The school has informed me that No. 65 has left the school and is now in the care of the Department of Community Services.

(Linda B, Field Notes, 14.03.12)

Even calling in the morning to check if students were at school was problematic because the distances to travel once we had confirmation that the students were present were very large. This is difficult enough during ordinary school hours (9 am to 3 pm), meanwhile behaviour schools operate on *reduced* school hours with a typical school day starting at 9 am and ending around 12.30–1 pm. This left a very short window of opportunity and, given that two-thirds of the population that we were working with had a history of severe learning and behavioural difficulties and that they took much longer to engage in the interviews and assessments than the mainstream/no behaviour group, the potential cost per student was extremely high if we could not achieve more than one student per trip.⁸ As schools often could not confirm attendance until sometime after 9.30 am, we had no choice but to get on the road and take what we could get.

⁸ 14 of 19 participating schools were located in Western or South-Western Sydney. Of these, 5 were located in South-Western Sydney, a 1.5 hr drive from base. A return trip to these schools therefore cost \$250 (including RA time, travel and tolls).

While our difficulties were often due to suspension and/or “jigging” of students in our two behaviour groups, we experienced a fair amount of difficulty with our mainstream/no behaviour group as well. For example, we arrived at School 6M for an appointment that had been confirmed the day prior to interview three students in mainstream with no history of disruptive behaviour (knowing already that another two students from our mainstream behaviour group were on suspension from School 6M) only to discover that 3 of the 4 students scheduled for that day were at a Year 7 orientation at their local high school. Given that School 6M was a fair distance from any of our other participating schools, the second author completed other work while waiting for the Year 6 s to return. We had hoped that she would be able to see the one remaining student but this was not to be:

Today I saw No. 40 at School 6M. After settling in and setting up, I rang the office to send the first child over. I had three year 6 children on the list plus No. 11. After a while they discovered that all the year sixes were at the high school for the morning session and Mrs M didn't actually know about this. So then we decided to try No. 11, but he had had a behaviour meltdown this morning. I said I was quite happy to try seeing [him]. Then we discovered that his mother had been called to come and pick him up so I didn't even get to start with him. Mrs M took me over to the “behaviour classroom” or whatever it is called just to see what was going on. No. 11 was there-he had calmed down and I'm sure he would have been fine doing at least a couple of language tests but his mum arrived about ten minutes later and he was gone for the day. Luckily, I had brought some other things to do, so I spent from 10 am to 12 noon being non-chargeable. The year 6 children were back for the afternoon and I saw No. 40 in two stages with lunch in between... And that was it for the day. We finished at 2.55 pm.

(Linda B, Field Notes, 08.11.11)

The difficulties in recruiting and retaining participants led to increased pressure when it came to ensuring that any face-to-face interaction resulted in useable data. This was not a certainty, given the young people that we were working with. Indeed, when the first author met with senior personnel within DEC and with each of the behaviour school principals to explain the research and invite their participation, one of the most common comments received was “So, how are you gonna get ‘em to talk?”

Angry, adolescent males are not known for being talkative, especially about their feelings. Knowing this, the decision was made to start each session with the assessments, leading with the Peabody Picture Vocabulary Test (PPVT), as this would allow time for our participants' to check us out and time for us to build rapport before moving on to more personal questions. We also ensured that the interview schedule was age-appropriate with short, linear, concrete questions that were easily understandable for a young person with limited patience and language (Graham 2010). Given that we wanted to foreground the young peoples' voices, the interview questions focused on issues that would be relevant to them and their schooling experiences; however, since there has been little research conducted with the target population and we could no longer afford to do school-based observations,

there was a lot that we needed to know from our participants. The final Round 1 interview schedule featured 8 sections with 64 base questions that would be asked of every participant. Those who had moved schools or who were enrolled in a behaviour school would be asked a further 21 questions. At least two implications flowed from this, the first being that we needed to engage young people with a reputation for being hard to engage, and second, we needed to elicit responses that we could use. Both challenges were expensive for different reasons. In the following section, we recall time spent with a number of our participants to showcase why.

Lollies, chess and Lego

No. 10 was a 13 year old “school refuser” whose mother called to request that we see him at home where she felt that he would feel safer talking about his school experiences. No. 10 lived in a public housing estate in a disadvantaged area of Sydney and, while he had been referred to School 4B for violence, we were advised by the school that this was mainly directed at school bullies and that it should be safe for adults to visit him at home. As university research safety policy requires that two researchers attend home visits to potentially dangerous locations, both authors met at the address provided. When we walked up the drive to the entrance of the little semi-detached villa, someone darted from one of the chairs outside the door into the neighbouring villa. The pungent aroma of marijuana greeted us at the doorway.

Undaunted, we rang the bell and No. 10’s mum opened the door. She was bubbly, friendly and completely lucid. She welcomed us into a small living room and called No. 10. It was dark inside the house with blinds covering the windows and it took some time for our eyes to adjust. The only light was emanating from a computer screen tucked away in the corner of the room. Hunched in front of it was a pale skinned, blonde haired boy who seemed oblivious to our presence. His mum asked if we would like a cup of tea while we set ourselves up on the dining table. She seemed anxious to talk and grateful that someone was interested in learning more about her child. No. 10 continued to ignore us.

Eventually his mother managed to coax him away from the computer and No. 10 approached the table. He was surprisingly tall once he uncurled himself from the little computer chair and very lean but well-built and strong looking. We cheerfully introduced ourselves, explained the research, asked if he consented to participate, and showed him the PPVT as an example of what we were going to do in the session. No. 10 nodded and signed the consent form but seemed barely conscious. When he did look at us, he did so by leaning his head back, so that he could see through half-open eyes. His speech was slow and slurred, which prompted his mother to explain that she’d “upped” his medication in anticipation of our visit. On the phone, she had told the first author that her son was on 5 concurrent medications: Seroquel (an antipsychotic), Endep (a tricyclic anti-depressant), Ritalin (a stimulant), Catapres (an anti-coagulant typically used to treat hypertension but also used off-label to treat ADHD), and sodium valproate (an anticonvulsant typically used to treat epilepsy but also used off-label for children with autism). These, she explained, had been prescribed to No. 10 when he had spent time in the Child and Adolescent Psychiatric Unit at the Children’s Hospital. His mum had explained on the phone

that No. 10 used to be violent towards her and that she had relinquished care to family services for 6 months when he was 11 years old but that they were both much better now.

No. 10 certainly seemed very dependent on his mother and was reluctant for her to leave the room. At the same time, he did not seem to want to answer any questions on the “How I feel?” screening form and, as such, we felt that his mother’s presence was both a help and a hindrance. We decided to focus purely on the language assessments in that session and moved back to the PPVT—a test that took most of our 13 year olds an average of 20 min to complete. No. 10 took an hour. This was not just because he appeared to be labouring under the influence of psychoactive medication. It was also because No. 10 was exacting; of himself, of us and the PPVT. Any sense of illogicality in the questions being directed towards him and he would stop to correct us, pointing to any potential for double-meaning. He also demonstrated fear of failure; refusing to answer any questions of which he was unsure of the answer. His pauses became so protracted that Linda B felt the need to explain that since the PPVT is multiple choice, a guess had a 1 in 4 chance of being right. But No. 10 said he would not guess. She then gently explained that no answer would have to be marked as wrong. No. 10 responded that at least that would not be *his* mistake.

Upon reaching a ceiling after an hour on the PPVT (in which No. 10 achieved a standard score of 106), we decided to call it a day and catch up with him again in the New Year.⁹ His mother told us that he was transferring to a support class in a mainstream school and, if that proved successful, she was happy for us to re-visit him there. Our second visit took place 3 months later, taking another 5¼ hours including travel time. Linda B relayed her experience later that day through her one of her now common “journal” emails; a practice that we had started to try and track the complexity of the research process but which ended up acting as a vital way to connect and to process what was happening:

I didn’t know what to expect from No. 10. He had grown taller and was more upright and lanky, but still pale. His hair was very short and he had quite a few little bites on his face. Mrs C [his teacher] said that he was quite happy to come and see me. Firstly, I showed him what he had already completed and showed him what else we had to do and said that he could have a break whenever he wanted... The interview went off on a couple of tangents when No. 10 was holding the floor on his favourite topic. There were a couple of awkward moments-I hope there is enough information there. It was sometimes difficult to follow what he was talking about and bring him back to the questions... About halfway through the interview, No. 10 lay down on the floor on his back then he curled up on his side and shut his eyes (I kept talking to him!) Then he went under the desk for a while. You’ll hear him showing me how he can do push-ups too, and also he stood up and measured how tall he was compared with me. Near the end of the interview No. 10 said that he

⁹ Trip 1 to see No. 10 took 4 hrs including travel time, time to set-up, build rapport, and trial which assessments he would do, plus 1 hr on the PPVT. Trip Cost: \$245.79. Trip 2 took 5.5 hrs and cost \$307.75. Neither of these costings take account of Author 1’s time which was contributed by the Administering Organisation.

knows that he is bad at everything and is screwed and doesn't care anymore. He also talked a bit about taking medication... After morning tea, he came back and sat on the chair and he was putting his feet up on the chair and the table sometimes. He was very keen for me to play chess with him although I told him I can't play. He got the chess game out, set it up and told me the rules in about two minutes. Unfortunately, it didn't help me much so he played on my behalf. I'm not sure who ended up winning but he kept telling me which piece could move where.

(Linda B, Field Notes, 27.03.12)

No. 10 was by no means an "outlier". Although he was the only scholar of chess in our cohort, we also had a keen origami expert who insisted on teaching his interviewer between questions, a few Lego enthusiasts (for one of whom we were advised to arrive with Lego in tow), and a tremendous number of fidgeters. To keep our young people interested and occupied, Linda B was armed with a large supply of "snakes" (of the no artificial colours or flavours variety) and, although Linda G was of the view that 10 snakes per participant should do it, we soon found that our young crew could chew through half a packet in a 50 min interview. Despite their penchant for words beginning with F and ending in K, most said "please" and "thank you" when they were offered their lollies and some even said "pardon" when they didn't fully hear an interview question.

Not surprisingly, many of the interview recordings were quite difficult for the transcribers to decipher. While No. 10 rolled on the ground and mumbled to himself, others spoke very softly and/or very quickly. Some, like No. 52, grunted or shrugged their way through the interview, while many more responded with the proverbial "I dunno". These, however, were anticipated difficulties and the experience, manner, and skill of the interviewer was always going to be of paramount importance. The ability to extract a usable answer from a naturally suspicious young person with limited language is not universal and interviewing these young people is an art. Friendliness can be perceived as being over-familiar and forward, relegating the interviewer to the perilous category of "try-hard", while professional distance and formality is associated with authority figures and can earn distrust. Persistence is irritating, especially when the problem is the question. And hubris is never tolerated. Listening to and respecting the young person—talking to them as an equal and not judging even when they say or do something inappropriate—go a long way towards establishing the necessary credibility and trust to make it through some 75 questions about things they don't like.

No. Nup. Nah... Nuh!

Fortunately, while some of our young people had few words, they were high on inflection. Accurate transcription became more important than ever because the way in which a word was said could change its meaning or its import. Unfortunately, the value of transcription is poorly understood in the clinical sciences and it is often one of the first budget items to be cut from grant applications, particularly when assessments are conducted by multidisciplinary panels. Text, however, is the lifeblood of qualitative researchers. While patterns and themes can emerge during the interview process or from listening to the audio recordings, this is rudimentary

analysis at best. Deep cross-theme analysis cannot be achieved by listening alone. Interview transcription enables researchers to juxtapose what they have heard with what they are reading, enabling them to draw new connections between groups of data and the development of higher-order insights. High quality and accurate interview transcription is also essential for the validity of the research process itself—whether qualitative or quantitative in design. In this mixed-methods project, for example, the professional transcribers were instructed to faithfully transcribe everything as it was said and not to correct grammar or pronunciation.

As not all transcribers were fully accurate and some would miss or misunderstand words, acronyms and phrases, the first author checked and corrected all 96 first round transcripts against the recordings. This was important for two reasons: first, because it is impossible to represent young people's voices when that voice is overlaid by the conventions of others, and second, because our young people had highly distinct ways of saying 'No' and when it came time to code the transcripts, it was clear that misinterpretation could skew the analysis and therefore, our results. For qualitative researchers, faithful representation of voice is a deeply ethical issue. It is not enough to simply read off what has been said in isolation, as might sometimes be the case in other disciplines, but to know the data well enough to judge whether what has been said at one point in time is being read and interpreted in a holistic sense.

For example, in the interview we asked our young people what they thought was the purpose of education: *What do you think school is for? What's the purpose?* Responses to this question were easy to code as most answers fell into two main categories: (1) to learn/get an education; or (2) to get a job. Our next question however, *Is that important to you?* raised an ethical question when No. 27 responded "No, not really". The codes for this question were: No: 0, Yes: 1, Ambivalent: 2, Don't know/No answer: 3. The two coders were split with one recording a zero for 'No' and the other recording a '2' for ambivalent. After some tense deliberation, the answer was eventually recoded with a '2' for ambivalent because the young person in question had no difficulty using a definitive 'No' elsewhere in the interview, whereas in this case, he used the qualifier 'not really'. For example in response to the questions, *Did you think that was fair?* and *Was it hard to make new friends?* he responded with an economical 'No'. Elsewhere when No. 27 thought further explanation or emphasis was necessary, he would qualify. For example, to the question *Can you remember any teachers that you had a really good relationship with?* he answered, "Nope. Really, really good, no."

Quality transcription enables electronic searching and a more holistic approach to data analysis. It also aides thoroughness when the answer to a question does not necessarily reflect the construct for which the question was designed. For example, in response to the question, *Is that important to you?* No. 96 also replied "Not really". Even though this was definitely an ambivalent response, inductive analysts would still treat this response cautiously because of his answer to the prior question, *What do you think school is for?* To this, No. 96, who already had an apprenticeship lined up for the following year, responded "To educate you to be smart"—an answer which changes the meaning of the structured question to follow. Simple numerical content analysis is based on the researcher's intended construct; in this

case, whether schooling is important to these young people. However, in response to the question *Is that important to you?* No. 96 is clearly saying that “to be smart” is not of most importance to him, rather gaining life skills is important, and that there is more to learn *beyond* the academic knowledge that is taught in schools.

- LB: What do you think school is for?
96: To educate you to be smart.
LB: Is that important to you?
96: Not really.
LB: No. Why not?
96: Just school doesn't teach you life skills.
You've got to get more of life than you've got to do at school.

Analysing such responses reliably and ethically is of utmost importance in qualitative research and it was for precisely such ambiguous responses that the “Ambivalent” code was created and an inductive approach to content analysis was chosen. While a simple, purely quantitative analysis could report the percentages of students who said that education was or wasn't important to them, the number or percentage of students who gave an ambivalent or equivocal response is far less important than the actual answers they gave. At this point, it is fitting to return to No. 27 whose response “No, not really” sparked this defence of interview data analysis. This short but compelling excerpt places his statement into context without which No. 27's words could be taken to mean that he, and students like him, do not care about education or their future.

- LB: Okay. What do you think the purpose of school is?
27: The purpose of school?
LB: Yes.
27: To learn.
LB: Is that important to you?
27: No, not really.
LB: No? Why not?
27: Learning, yes it is but not much.
LB: Not so much?
27: No because I've got other things to worry about.
LB: What other things do you have to worry about?
27: My mum dying and all that kind of stuff.
LB: Is your mum dying?
27: No but I worry that she will.
LB: Why are you worried about that?
27: I don't know. If my mum dies I won't be coming to school any more.
LB: Is your mum sick?
27: No, she ain't sick. It's just I worry about her. She's under lots of stress.

- LB: Is she? What sort of stress is your mum in?
 27: Very hard stress cause she's getting cramps in her arm and I'm very worried.
 LB: Are you?
 27: Cause my mum's the only thing close to me really. If I lost her then I, well I'd be just...

Popular discourse already positions these young people as “menaces” who do not value education (see McDougall 2011) and it would be unconscionable if our research was to perpetuate that perception because we had failed to take the time and make the effort to read beyond singular questions and answers. Whether she was dying or not, No. 27 is clearly worried about his mother—as were a number of participants in our two behaviour groups. Not only did their troubled home lives overshadow the significance of school but a perceived lack of empathy on the part of teachers and principals was a source of great resentment. So too was a perceived lack of respect for their parents and family. No matter how dysfunctional the family or their relationship, our young people were unswervingly loyal to their main caregiver; this was usually their mothers. Even though he now lives with his father, this is evident in the way that No. 22 spoke about his troubled mother, whom he still sees and clearly loves:

- LB: Why did you move school this time?
 22: From there I went to DOCS. I went into DOCS. From [that primary school] I went straight into DOCS.
 LB: DOCS. Yeah, why did you go there?
 22: Um, because my mum, she was a real bad alcoholic, then. But, now she's different. She used to drink so much that she end up start spewin' blood, but she keep on drinkin'. But, now she stopped and she used to be on weed, like, marijuana and yeah, that's it. She always used to smoke marijuana and that.

While No. 22 describes behaviours that are often used to criticise parents of disadvantaged students and that resulted in his removal from his mother's care, his mother is still his mother and he still loves and needs her. This became clear when we asked No. 22 to describe his personality, at which point No. 22 again referred to his mother to explain how and why he was “protective”:

- 22: Doesn't matter what you say to me as long as you don't say nothin' about my mum. Somebody says somethin' about my mum, the next day...that day, they'll die. They'll die for sure. 'Cause nobody....I swear on all my whole dead ancestor, nobody says nothin' about my mum. Last time somebody said somethin' about my mum. That was it. They almost went to hospital. But, instead, I went to hospital because I popped my knuckle again. But, I didn't care. I was all fired up.

Doing research with sad and angry young people is harrowing. It is very difficult to listen to their stories and to not be able to do anything practical or immediate to help. Simply bearing witness to those stories—without trying to rationalise or reinterpret—did seem to be appreciated by some of our participants. These stories however become a part of the listener's own life as intensely private moments shared between two strangers are impossible to forget.

In the ARC proposal, the first author made a point of arguing for an experienced and highly-skilled research assistant. Despite the large funding shortfall, the decision was made to appoint and retain Linda B, rather than a “cheaper” and much less experienced graduate. This was a decision that proved critical to the research; particularly with respect to our longitudinal behaviour group who would not take kindly to a chorus of inexperienced and/or indifferent casuals. Despite our collective experience however, we still found it disturbing to see a pale 13 year old boy living in a medicated bubble with the blinds closed and X-Box Live as his only connection to the outside world. We also found it hard to listen to boys of 10 and 11 years of age tell us how they protected their mothers from junkies and how they were scared that their mothers would die if they kept drinking or their partners kept bashing them.

After such interviews, we would debrief on the way home to try and process what we had heard, however, as most of the interviews were conducted by Linda B, we developed a system where we would touch base by phone either during or at the end of a long day. These costs are seldom acknowledged and rarely budgeted for, despite their existence and their financial implications. Close contact with committed research assistants however enables researchers to remain intimately connected with their own project, even if they might live some of it second-hand. The question for our discipline in the current research climate is how to write such costs into an ARC proposal without the project being dismissed as second-rate, particularly when it comes to arguing for funding in a field of research that is already poorly understood and too often undervalued.

Conclusion

While all researchers are grateful to receive public funding for their projects, particularly Early Career Researchers, the underfunding of research presents ethical and practical problems that have real-world implications. This is particularly the case when it comes to doing research in difficult contexts with difficult young people. Research in schools is messy business. They are often chaotic places with agendas and timelines that do not gel well with academic research designs; that is, the type that is likely to be successful in an ARC application. Disaffected young people are even less accommodating than their schools. But the difficulty of doing this type of research only increases its value. Given that an enrolment place in a behaviour school costs taxpayers between 3 and 10 times that of a place in the local secondary school, the importance of working with and learning from young people who have been referred to behaviour schools is surpassed only by the importance of tracking down and speaking with the ghosts who refuse to attend. Once that happens

however there is the economic cost of ensuring that what they say is accurately represented and the emotional cost of honouring the sadness and hurt from where those words come.

Unfortunately, these factors are not well understood by our peers in the human sciences and there remains a common perception that education research lacks rigour, particularly qualitative approaches. This is problematic, given that this is the pool from which our ACG assessors are drawn. It is well known that scientists have worked hard over the last few decades to communicate the value of research in the clinical and natural sciences and that they have been successful in raising the profile and prestige of scientific research. Given the contraction in education research funding in recent years and the apparently poor performance of Division 13 (Education) in ERA 2010, it is now critical that researchers in education speak up—not only about the value of the work they do—but about the beauty and complexity of research in this field, the critical role that qualitative approaches to data collection and analysis play in ensuring quality, and the insights that this type of research is able to produce.

Acknowledgments This research was supported under the Australian Research Council's *Discovery Projects* funding scheme (DP110103093). The views expressed herein are those of the author and are not necessarily those of the Australian Research Council. Ethics approval was obtained from the Macquarie University Ethics Committee (Final Approval No. 5201000654) and the NSW Department of Education and Communities (SERAP No. 2011027).

References

- Achenbach, T. M. (1991). *Integrative guide to the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington: University of Vermont, Department of Psychology.
- Allen Consulting Group (2008). *Recognising the full costs of university research*. Report to the Department of Innovation, Industry Science and Research, Canberra.
- Bazeley, P. (2003). Defining 'early career' in research. *Higher Education*, 45, 257–279.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd ed.). Thousand Oaks: Sage Publications.
- Dunn, L. M., & Dunn, L. M. (2007). *Peabody picture vocabulary test* (3rd ed.). Circle Pines: American Guidance Service.
- Goodyear, P. (2008). Educational research and the ERA, *AARE News, Australian Association for Research in Education (AARE)*, 63, p. 4. Available at: <http://www.aare.edu.au/news/newsplus/news63.pdf>.
- Graham, L. J. (2010). Thinking pedagogically, Chapter 11. In L. J. Graham (Ed.), *(De)Constructing ADHD: Critical guidance for teachers and teacher educators* (pp. 205–220). New York: Peter Lang Publishing.
- Graham, L. J. (2012). Disproportionate over-representation of Indigenous students in New South Wales government special schools. *Cambridge Journal of Education*, 41(4), 163–176.
- Graham, L. J., & Sweller, N. (2011). The inclusion lottery: who's in and who's out? Tracking inclusion and exclusion in New South Wales government schools. *International Journal of Inclusive Education*, 15(1), 941–953.
- Graham, L. J., Sweller, N., & Van Bergen, P. (2010). Detaining the usual suspects: charting the use of segregated settings in New South Wales government schools, Australia. *Contemporary Issues in Early Childhood*, 11(3), 234–248.
- Graham, L.J., Van Bergen, P., & Sweller, N. (2011). *Tracking the experiences of students enrolled in segregated settings for challenging behaviour and their reintegration to mainstream schools*. Australian Research Council (ARC) Discovery Project: DP110103093.

- Granite, E., & Graham, L. J. (2012). Remove, rehabilitate, return? The use and effectiveness of behaviour schools in New South Wales, Australia. *International Journal on School Disaffection*, 9(1), 39–50.
- Healy, G. (2009). Better rate, less cash. *The Australian*, Higher Education Supplement, October 28. Available at: <http://www.theaustralian.com.au/higher-education/better-rate-less-cash-in-arc-grants/story-e6frgcjx-1225791901778>.
- Marginson, S. (2007). *Are neo-liberal reforms friendly to academic freedom and creativity?* Centre for the Study of Higher Education, The University of Melbourne Seminar, Ideas and Issues in Higher Education, May 28.
- Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist*, 63(3), 160.
- McDougall, B. (2011). *Too bad for school—scared principals refuse to take back dangerous principals*. The Daily Telegraph, June 3rd. Available from: <http://www.dailytelegraph.com.au/news/sydney-news/too-bad-for-school-scared-principals-refuse-to-take-back-dangerous-pupils/story-e6freuzi-1226068259625>.
- Peters, M. A., & Besley, T. (2006). *Building knowledge cultures: Education and development in the age of knowledge capitalism*. Maryland: Rowman and Littlefield.
- Schools Australia, 4221.0 (2011). Australian Bureau of Statistics: Canberra.
- Spielberger, C. D. (1973). *State-trait anxiety inventory for children: Preliminary manual*. Palo Alto: Consulting Psychologists Press.
- Takkashori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Williams, K. T. (2007). *Expressive vocabulary test* (2nd ed.). Circle Pines, MN: Pearson.
- Yin, R. K. (2003). *Case study research: Design and methods*. (3rd Ed., Vol. 5). Thousand Oaks, CA: Sage Publications.

Author Biographies

Associate Professor Linda J. Graham is Principal Research Fellow in the School of Cultural and Professional Learning, one of three schools in the Faculty of Education at Queensland University of Technology (QUT). She is Lead Chief Investigator on ARC Discovery DP110103093 which she shares with Dr Penny Van Bergen and Dr Naomi Sweller from Macquarie University. Her work focuses on institutional contributions to student disaffection and the improvement of responses to children who are difficult to teach.

Mrs Linda Buckley is a sessional academic in the Department of Linguistics, Macquarie University. She has considerable experience working with challenging young people and was responsible for the majority of the data collection on this project. Her journal notes were conducted in her own time and formed the basis of her contribution to the writing of this paper.