Cognitive mechanisms for the formation of public perception about national testing: A case of NAPLAN in Australia

Jihyun Lee1 · Jung-Sook Lee2 · Jodi Lawton1

Received: 4 January 2022 / Accepted: 19 August 2022 / Published online: 5 September 2022
© The Author(s) 2022

Abstract
Although there has been intense criticism of NAPLAN in educational policy debates in Australia, little scholarly efforts have been made to understand the underlying cognitive mechanisms that contribute to the public narrative about the national testing program. We aim to provide tentative evidence about the way public perceptions about NAPLAN may be formed. Our results show empirical support for the incentive, interpretative, and institutional effects, which suggest ways that national testing program can be improved. That is, it needs to (a) provide a diverse range of incentives to promote people’s self-interest (incentive effect); (b) demonstrate good alignment with the core values, social norms, and attitudes of the given society (interpretative effect); and (c) build a consensus about the institutional use of the test results (institutional effect). We conclude with practical implications and recommendations about seeking public support for the seemingly unpopular national educational policy.

Keywords NAPLAN · National testing · Public opinion · Incentive effect · Interpretative effect · Institutional effect

“A brilliant system but being practiced and reflected in the worst means. While the intention is good, its usefulness is very much questionable, and it is becoming manipulative” [Participant #51]

Over the past decade in Australia, no educational initiative has created more heated discussion than the National Assessment Program – Literacy and Numeracy (NAPLAN). NAPLAN was implemented in 2008 as a census-type testing to annually assess and monitor literacy (reading, writing, and language conventions) and numeracy skills of all students in Years 3, 5, 7, and 9. The main purposes of this
national testing are known as (a) reporting on student learning growth, (b) guiding school improvement, (c) strengthening system accountability, (d) informing parents/carers about student and school performance, and (e) monitoring national, state and territory programs and policies (McGaw et al., 2020). In spite of these ambitious goals and strong support from the successive governments of both major parties of Australia, NAPLAN has been heavily criticised in media outlets and publications including the print media (e.g. Reid, 2020), academic book chapters (e.g. Cumming et al., 2016), journal articles (e.g. Hardy, 2014; Klenowski & Wyatt-Smith, 2012), trade journals (Zadkovich, 2017), and reports commissioned by various teacher unions (e.g. Canvass Strategic Opinion Research, 2013). Needless to say, significant news media coverage on NAPLAN has exposed many Australians with stories about its negative consequences on the education system and teaching profession (Shine, 2015; Shine & Rogers, 2021).

The present study focuses on one specific aspect of NAPLAN, i.e. public perceptions about NAPLAN. Particularly, we aim to provide some, admittedly tentative, understanding of how members of the general public may form their perceptions about a mandatory national testing program such as NAPLAN. In the educational assessment literature, the term ‘public’ is defined as “an umbrella characterization of multiple stakeholder groups with an interest in assessment results” (Buckendahl, 2016, p. 454). Further, ‘public opinion’ is defined as “opinions concerning social or governmental matters rather than private matters…. … It comprises opinions, attitudes, preferences, beliefs, and values” (Pizmony-Levy & Bjorklund, 2018, p. 240). Members of the general public often hold opinions about social or governmental matters, but it becomes public opinion when their opinions are communicated to other members of society through media discussions or scientific studies. On the other hand, there is a narrower definition of public opinion, as Perrin and McFarland (2011) claimed. That is, individuals’ opinions become public when opinions are gathered through technical practices such as formal conduct of surveying a representative sample of the population. For the purpose of the present study, we use the term ‘public’ to simply refer to members of the general public. Following on this, ‘public opinion’ or ‘public perception’ in this study means the opinion or perception held by members of the general public regarding the issues that are generally considered to be in the public domains such as education, health, or social security. ‘Members of the general public’ in this study means that they are not experts on the topic in question, nor had substantial personal or professional experiences related to it.

The present study seeks to contribute to our understanding of how members of the general public may respond to national testing the way they do. As mentioned above, in spite of a plethora of studies delineating specific accounts regarding the negative consequences of NAPLAN (e.g. see a systematic review by Rose et al., 2020), there has been little scholarly attention given to examining cognitive mechanisms that facilitate the formation of the public’s perceptions about the national testing program. Cognitive mechanisms allow people to receive information and interpret and evaluate and sometimes act upon that information (Heinström, 2010; MacKay, 1969). The present study reflects on theories and research in the area of public opinion formation and change in social policies (see Busemeyer, 2012; Busemeyer & Garritzmann, 2017; Pierson, 1993; Jacobs & Weaver, 2015) and studies
on public opinion on standardised national/international assessments (e.g. Jacobsen et al., 2013, 2014; Pizmony-Levy et al., 2018), and explores cognitive mechanisms underlying the perceptions of NAPLAN held by members of the general public (i.e. New South Wales residents) in Australia.

We believe it is important to examine public perceptions regarding educational matters including national testing. This is mainly because the government’s educational policies, including national testing programs, cannot be sustained or thrive without public support (Berkman & Plutzer, 2005; Jacobsen et al., 2013, 2014; McDonnell, 2000; Pizmony-Levy & Bjorklund, 2018). Social and educational policy scholars have elaborated ways in which this relationship between public support and public program holds for successful implementation and sustainability of national educational programs. First, a national testing program may often undergo regular review or need to implement necessary changes and innovations. For these types of government activities, public funding is required, and public perception is critical to secure support for the program funding (Jacobsen et al., 2013; Lewis & Hardy, 2015; McDonnell, 2000). Secondly, it is hard, if not impossible, to reflect on the expectations from members of the public without understanding their views on the national testing. The ultimate goal of a national testing program is to inform all stakeholders about the effectiveness of the educational system and to gather information about how to serve them in a way that the majority is satisfied (Jacobsen et al., 2013, 2014; Pizmony-Levy & Bjorklund, 2018). The identification of the possible sources of negative feedback and understanding of how the public may feel about national testing is critical to judge what action may be required to address identified deficiencies and to enhance the public’s sustained willingness to engage with the program.

Thirdly, when the public forms and expresses strong opposition to a particular educational program, they call for changes or request the primary actors to be accountable for perceived failures (Jacobsen et al., 2013; McDonnell, 2000). In this sense, the survival of an educational policy, including a national testing program, largely depends on public satisfaction with the program. Finally, scholarly efforts to theorise cognitive mechanisms whereby members of the public form their perception about national testing is important because such information can allow for predictions of public perceptions about future national educational programs. Such information can also enlighten broader implications such as public satisfaction with school quality (Jacobson et al., 2013) or the education system as a whole (Pizmony-Levy & Bjorklund, 2018).

1 Criticisms about national standardised testing

In the mid-1990s, international large-scale assessments such as the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) started to gain global attention as the international assessment results sparked serious conversations about the success or failings of educational systems around the world (Fischman et al., 2019; Sellar & Lingard, 2013). What followed globally, and in particular among Anglo-Western countries
such as Australia, Canada, the UK, and the USA, was the uptake in the governments’ approach to strengthening the educational accountability system by using national standardised testing results as evidence for school effectiveness (Hargreaves, 2020; Lingard & Sellar, 2016; Sellar & Lingard, 2013). This approach has gained its momentum since the late 1990s, spurred by the idea that educational systems would improve by applying business techniques and market principles to school systems to hold the underperforming systems or schools accountable (Lingard & Sellar, 2016).

In spite of the intention of making failing schools/systems accountable, scholars have expressed deep concerns about using the standardised testing as the measure of school effectiveness. This approach was criticised for giving too much weight to the national policy on the school accountability measures driven by narrowly defining educational success (Hargreaves & Goodson, 2006; Ravitch, 2016). For instance, Canadian researchers (e.g. Hargreaves, 2020; Kearns, 2011) pointed out that standardised testing has failed to live up to the hype surrounding it. While promoted as encouraging well-being and equity through the accountability system, standardised testing has done the opposite as studies suggest the increase in the perceived feelings of shame and marginalisation among schools and students in disadvantaged backgrounds (Hargreaves, 2020; Kearns, 2011).

A similar sentiment is expressed among US-based researchers who argued that the school accountability system only adds the privilege to the already privileged while stigmatising students who do not learn well in a traditional way (e.g. Ravitch, 2016). It is known that a range of non-instructional variances directly influence standardised test results, such as the number of parents at home, poverty, and parents’ education levels. These variances need to be accounted for when examining the differences in the standardised testing scores in the USA (Kohn, 2000). Furthermore, with the heavy emphasis on the demonstration of the improvement, schools and teachers have increasingly relied on instructional styles that trivialise the learning process, narrow the school curricula, encourage teaching to the test, stifle creativity by emphasising identifying only one correct answer, and waste valuable teaching time for superficial learning (Berliner, 2011; Herman & Golan, 1993). Pressure to demonstrate growth in the performance measure also led teachers to engage in inappropriate test preparation and cheat when administering a test in this era of strengthening the standardised testing in England (Mattei, 2012). Decreased instructional time on non-tested subjects such as Arts and fitness is another aspect that is repeatedly recognised and criticised in both the UK and the USA (Kelly et al., 2018; Mattei, 2012).

Criticisms surrounding the use of national testing noted by US, UK and Canadian researchers have been similarly recognised by Australian researchers with respect to unintended consequences of NAPLAN. Researchers lamented about NAPLAN becoming the sole, standard measure of school effectiveness (Lewis & Hardy, 2015) while it continues to create the classroom culture of teaching to the test and narrowing the curriculum and pedagogic choices, encouraging practices that focus only on measurable outcomes, and dumbing down of learning (e.g. Brady, 2013; Harris et al., 2013; Lewis & Hardy, 2015; Ragusa & Bousfield, 2017). NAPLAN is also seen as marginalising students in disadvantaged backgrounds, increasing stress and anxiety among students, teachers, and parents (e.g. Brady, 2013; Canvass Strategic Opinion Research, 2013).
and rewording lower-order skills, and stifling creativity (Harris et al., 2013; Ragusa & Bousfield, 2017). There is a growing concern about the systematic misuse of data while contributing to high staff turnover and low morale (Ragusa & Bousfield, 2017). As can be seen, the nature and content of the criticisms about the negative impact of national testing is quite similar in the context of different countries.

It should be noted, however, that not all NAPLAN-related discussions have been negative. Recent reports (Cumming et al., 2018; Louden, 2019; McGaw et al., 2020) suggest that school sectors and teaching professionals were ambivalent and acknowledged both positive and negative consequences of NAPLAN including its contribution to guiding school improvement and meeting specific learning goal targets. A recent comprehensive review, ‘NAPLAN Reporting Review’ (Louden, 2019), prepared for the Council of Australian Governments (COAG) by Education Services Australia, noted that while the media reporting of NAPLAN has been predominantly negative, there has been some shift among the public and stakeholders about its value over time. Even the harshest critics of NAPLAN, the teachers’ union, acknowledge the importance of having a national testing program to support communities most in need, track the educational performance of all students, and monitor individual students’ progress over time. Parents/carers in particular perceive that the information about their children’s performance is more in-depth when it is being monitored over time and compared to the averages (national and school, year-level cohorts) in an external independent assessment (Ragusa & Bousfield, 2017). Although school system and sector authorities tended to express strong opposition to NAPLAN, they have also started to accept NAPLAN as part of Australia’s accountability system (McGaw et al., 2020; Ragusa & Bousfield, 2017).

2 Theories about the formation of public perception

We have adopted a theoretical framework to explain how public perception about the national testing such as NAPLAN may be formed. The central idea of the framework was borrowed from the research area of “policy feedback” (Jacobs & Weaver, 2015; Pierson, 1993). It suggests that “policies produce politics” in that “public policies were not only outputs of but important inputs into the political process, often dramatically reshaping social, economic, and political conditions” (Pierson, 1993, p. 288). The framework that we adopted consists of three main concepts, known as incentive, interpretative, and institutional effects (see Busemeyer, 2012; Busemeyer & Garritzmann, 2017). In the following section, we elaborate on each of the three concepts in the context of NAPLAN.

2.1 Incentive effect

The incentive effect (see Busemeyer, 2012; Busemeyer & Garritzmann, 2017; Pierson, 1993) postulates that self-interest and personal experiences may play the most critical role in forming people’s perceptions about public policies. It has been argued that “the ways in which political systems confer resources on individuals and
create incentives for them are the bread and butter of contemporary political science” (Pierson, 1993, p. 598). Thus, it hinges upon the argument that people will naturally support social and educational policies that generate direct benefits for themselves. Many media and scholarly reports have claimed that the Australian public and stakeholders are not convinced about the benefits, incentives, and resources that NAPLAN can provide (e.g. Hardy, 2014; Ragusa & Bousfield, 2017; Zadkovich, 2017) although some positive shifts in these views have been noted (Cumming et al., 2018; Louden, 2019; McGaw et al., 2020).

Nevertheless, it still appears that NAPLAN does not meet the expectation of its primary stakeholders (e.g. schools, teachers, and students), which tends to generate disappointment, particularly about its diagnostic capability to meet the diverse needs of individual students. There is a legitimate reason for this expectation as it is claimed that national testing would assist student learning and guide school improvement (McGaw et al., 2020). On the other hand, many teachers and schools doubt NAPLAN’s capacity to provide any diagnostic or new information about their students beyond what they already know from in-school assessment data (e.g. Brady, 2013; Hardy, 2014; Lee et al., 2019). We would argue, though, that members of the general public may hold different views from those held by primary stakeholders (such as school-level actors). This could be the case, in particular, if parents and other concerned citizens perceive national testing as a central instrument to identify ‘failing’ schools and teachers and to provide evidence suggesting the needs for system-level support and direction (Jacobsen et al., 2013, 2014). In this way, members of the public may perceive national testing useful in ensuring the evidence-based educational accountability system (Lewis & Hardy, 2015).

Social and educational policy scholars have long argued that public discussion surrounding educational matters occurs not only among those who are directly affected such as teachers, students, and parents, but also members of the general public including taxpayers and employers (e.g. Jacobsen et al., 2013). Thus, the self-incentives argument can be applied to those who are primarily and directly affected by the policy (e.g. schools and teachers) as well as members of the public who may not be directly affected by the policy. In sum, the self-incentives argument dictates that the policy might be supported if members of the public see its benefit. Alternatively, the policy will lose a rationale for its existence if members of the public fail to perceive benefits from the program. Thus, we ask the first research question of this study: Would self-incentives explain perceptions about NAPLAN held by members of the general public?

### 2.2 Interpretive effect

While the incentive effect focuses on individuals’ concerns for their own self-interest and benefits, the interpretive effect points to group norms and social values as the determinants of the public’s policy preferences and attitudes (Busemeyer & Garritzmann, 2017; Pierson, 1993). Its main concerns are social norms, cohesion and justice, altruistic motivation, and consequences for the society, which tend to
be influenced by ideological and political or religious beliefs held by individuals or
groups (e.g. Busemeyer & Garritzmann, 2017; Pierson, 1993).

The [Australia’s] national declarations, commencing with the 1989 Hobart Decla-
ration on Schooling (MCEETYA, 1989) and reiterated in the 2008 Melbourne
Declaration on Educational Goals for Young Australians (MCEETYA, 2008), artic-
ulate the core educational values and central role of education in Australia. These
Declarations define the national goals to be the establishment of and support for
a democratic, equitable and just society, with the dual purposes of excellence and
equity (MCEETYA, 2008). To achieve these aims, provision of learning support for
individual students with diverse needs and backgrounds is within the responsi-
bility of schools/teachers (Cumming et al., 2018; Wyatt-Smith & Klenowski, 2010).
The significant value placed on individualised learning support is more of a norm
than an exception in many Australian classrooms (Mills et al., 2014), as evidenced
in various forms of differentiated instruction and pedagogical approaches to cater
to the diverse needs of students. Indeed, NAPLAN was developed based on the
principle of equality, i.e. to ensure monitoring and provision of quality education
to all Australian students (Cumming et al., 2018) and to enable all students to have
equity of opportunity for success in life (McGaw et al., 2020). It is noted that “there
remains… a public interest [by the community, parents/carers, schools, and other
members of the public] in making available information about each school’s con-
tribution to the national effort to make Australia a high equity, high performance
nation” (McGaw et al., 2020, p. 133). Thus, the second research question, based
on the interpretative effect, asks *Would people’s perception of NAPLAN reflect the
core educational values of Australia that emphasise the needs of individual students,
social solidarities, and equity?*

2.3 Institutional effect

Finally, the institutional effect points to the important role of the institution respon-
sible for the development of and communication about the program. In a broad sense,
the institution can be viewed as a formal organisation (typically government) that
aims to deliver a public good to members of the society (such as education, health,
or social security). Education is generally considered to be a public good in a soci-
ety where public education is run by a democratically run government (Levin, 1987;
McDonnell, 2000; Musgrave, 1959; Ragusa & Bousfield, 2017). This means that the
government is the primary actor to run the education system and that members of
the general public pay taxes with the hope that their money will be effectively used
by policy makers (Jacobsen et al., 2013; McDonnell, 2000). The institution, for the
purpose of the present study, is defined as the government sector responsible for the
development and management of the national testing program. The institution that
designs, develops, and implements the policy often also disseminates, manages, and
controls the information flow to the general public (Pierson, 1993). In this process,
the public’s interpretation of the policy is inevitably influenced by how the institution
organises and disseminates its policy (e.g. Pierson, 1993; Svallfors, 2012). Therefore,
what motivates members of the general public to express their opinions may not be
just about the policy event itself (e.g. NAPLAN testing) but the way the event and its outcomes are organised, advertised, used, and communicated by the institution.

The policy feedback literature explicates that the institutional effect can be augmented by the visibility and traceability of a particular policy measure (Pierson, 1993). That is if policy outcome is manifested like a “slow drip” (e.g. the impact of policy change in the pension system or health care system), it is less likely to mobilise immediate reaction following new policy implementation. On the other hand, if a policy outcome is displayed in a “single package” (e.g. transportation failure), the event outcome becomes highly visible often in the media, which can attract the mass public’s interest and immediate reactions (Jacobs & Weaver, 2015; Pierson, 1993). NAPLAN is a highly visible educational event in Australia. It is regularly and annually administered to all Australian students at the designated year-levels (i.e. population-based assessment). The testing unfolds regularly over several months and in a specific timeline, from May (testing month) to August/September (reporting to schools and individual students), followed by the national and state/jurisdiction reports in February the next year. Each of these time points attracts media attention and experts’ commentaries featured in the newspapers and on social media. The event outcomes (i.e. the test results) are also highly visible in a range of reporting channels, through reporting for individual students; the national and state/jurisdiction report cards; and most notoriously, the school performance results in the public domain through the website of My School that allows direct comparison of school performance among those with similar socio-economic profiles. Thus, it is fair to say that NAPLAN’s visibility is not limited to the central stakeholders (students, parents, teachers, and other school-level actors) but reaches to other members of the public who may not be directly affected by the national testing.

In addition to its highly visible nature, NAPLAN has a high level of traceability (i.e. information is traceable in terms of its owner and the history of the assessment). The policy feedback literature suggests that the public’s opinion of a policy measure will be stronger when it is relatively easy to trace the main actor of the policy program (Pierson, 1993). NAPLAN is a federal government program, which makes it convenient for the public to comment on it as the government’s failure or success. Following on from this, the third research question of the current study is Would the program’s high visibility (due to the public nature of the event outcomes) and traceability (directed to the government), that are primarily created by the institution, contribute to the public’s interest and perception about NAPLAN?

In summary, the three theoretical assumptions presented above may provide a useful framework to explain cognitive mechanisms underlying the formation of perceptions about NAPLAN held by members of the public. To gain public support for a national program, it is necessary to understand what the public thinks about it, how the public may see its benefits and areas for improvement, and necessary revision strategies to maintain the public support for the program. Given the international context of growing scepticism about large-scale testing (Elliott et al., 2019), the information from this study may be used for management and future development of a national testing program not only in Australia but more broadly among the countries where there is a growing concern and public discontent towards the school accountability policy of standardised testing (Pizmony-Levy et al., 2018).
3 Methods

3.1 Data and participants

The present study is based on two rounds of data collection. The first dataset [Phase 1] was based on a convenience sample of university staff, consisting of both administrative/professional and academic staff working at the authors’ institution (according to the university website, there are 6799 full-time equivalent staff working at this university). We sent out an online survey link that contained a free-format response question, “What is your view of NAPLAN?” The survey link was open for about four weeks in April 2019.

Table 1 presents sociodemographic composition of the sample of this study. As can be seen under the columns of Phase 1 ($N=89$), there were more females (48%) than males (20%) although 21% did not report on gender. Their age ranged from 18 to 59 overall, with more than 60% under age 49. Many of them (45%) reported having no child and only 17% reported having school-aged children. In this sample, 19% reported having a Bachelor’s degree, 16% a Master’s degree, and 37% a PhD. There was a spread of the income levels ranging from 10% reporting receiving less than $1000 fortnightly to 6% receiving $4000 or more fortnightly. Given the information on educational and income level, the majority of Phase 1 sample appeared to be administrative/professional staff rather than academic staff.

The second set of data collection [Phase 2] was conducted with members of the general public living in New South Wales (NSW). We employed a data collection company Qualtrics to reach out to potential participants. An online survey that contained a free-response question to write anything about NAPLAN was open for eight weeks from late October to early December in 2019. The sociodemographic composition of this sample ($N=62$) is illustrated in Table 1. Compared to the Phase 1 sample, the majority of this group was older than 40 years old (74%) and there was a better gender balance (58% females and 40% males). The majority of them (57%) did not have tertiary education, only 21% reported having a Bachelor’s degree, and nobody reported having a PhD. Among those who provided the answer on parental status, all of them reported having children with 26% reporting to have school-aged children (compared to only 17% of Phase 1 participants). There were 16% of regional residents and 24% of rural residents.

Overall, it appears that there were many differences in the sociodemographic backgrounds of Phase 1 and 2 participants. This does not mean that the combined data can achieve the representativeness of the NSW population, but it certainly suggests that the overall sample had members of the public with a diverse range of sociodemographic backgrounds. Appendix Table 2 presents detailed sociodemographic information about each participant in this study. The participant numbers from Appendix Table 2 are included at the end of each excerpt.

3.2 Analysis strategies: thematic analysis

We have employed thematic analysis. It is one of the most frequently used qualitative interpretation methods (Guest et al., 2012). The principles of thematic analysis illustrated in Braun and Clarke (2006) were applied to our data, which suggests the
Table 1  Sociodemographic background of the study participants

|                        | Phase 1  | Phase 2  | Total     |
|------------------------|----------|----------|-----------|
|                        | (N=89)   | (N=62)   | (N=151)   |
|                        | n        | %        | n         | %          | n         | %          |
| Age                    |          |          |           |            |           |            |
| 18–29                  | 18       | 20.2     | 8         | 12.9       | 26        | 17.2       |
| 30–39                  | 21       | 23.6     | 8         | 12.9       | 29        | 19.2       |
| 40–49                  | 18       | 20.2     | 15        | 24.2       | 33        | 21.9       |
| 50–59                  | 7        | 7.9      | 13        | 21.0       | 20        | 13.2       |
| 60 or older            | 0        | 0.0      | 18        | 29.0       | 18        | 11.9       |
| Non-response            | 25       | 28.1     | 0         | 0.0        | 25        | 16.6       |
| Gender                 |          |          |           |            |           |            |
| Female                 | 48       | 53.9     | 36        | 58.1       | 84        | 55.6       |
| Male                   | 20       | 22.5     | 25        | 40.3       | 45        | 29.8       |
| Non-binary/other       | 0        | 0.0      | 1         | 1.6        | 1         | 0.7        |
| Non-response            | 21       | 23.6     | 0         | 0.0        | 21        | 13.9       |
| Education attainment   |          |          |           |            |           |            |
| Year 10–12             | 2        | 2.2      | 14        | 22.6       | 16        | 10.6       |
| Vocational certificate | 0        | 0.0      | 11        | 17.7       | 11        | 7.3        |
| Diploma                | 0        | 0.0      | 10        | 16.1       | 10        | 6.6        |
| Bachelors              | 17       | 19.1     | 13        | 21.0       | 30        | 19.9       |
| Graduate diploma or certificate | 0 | 0.0 | 6 | 9.7 | 6 | 4.0 |
| Masters                | 14       | 15.7     | 7         | 11.3       | 21        | 13.9       |
| PhD                    | 33       | 37.1     | 0         | 0.0        | 33        | 21.9       |
| Other                  | 1        | 1.1      | 1         | 1.6        | 2         | 1.3        |
| Non-response            | 22       | 24.7     | 0         | 0.0        | 22        | 14.6       |
| Income level†          |          |          |           |            |           |            |
| <$1,000                | 9        | 10.1     | 12        | 19.4       | 21        | 13.9       |
| $1000–$1499            | 5        | 5.6      | 17        | 27.4       | 22        | 14.6       |
| $1500–$1999            | 12       | 13.5     | 17        | 27.4       | 29        | 19.2       |
| $2000–$2499            | 15       | 16.9     | 6         | 9.7        | 21        | 13.9       |
| $2500–$2999            | 7        | 7.9      | 3         | 4.8        | 10        | 6.6        |
| $3000–$3999            | 11       | 12.4     | 4         | 6.5        | 15        | 9.9        |
| >$4000                 | 5        | 5.6      | 3         | 4.8        | 8         | 5.3        |
| Non-response            | 25       | 28.1     | 0         | 0.0        | 25        | 16.6       |
| Employment status      |          |          |           |            |           |            |
| Full-time              | 50       | 56.2     | 17        | 27.4       | 67        | 44.4       |
| Part-time              | 13       | 14.6     | 14        | 22.6       | 27        | 17.9       |
| Unemployed             | 1        | 1.1      | 3         | 4.8        | 4         | 2.6        |
| Not in labour force    | 3        | 3.4      | 27        | 43.5       | 30        | 19.9       |
| Non-response            | 22       | 24.7     | 1         | 1.6        | 23        | 15.2       |
| Occupation             |          |          |           |            |           |            |
| Manager                | 0        | 0.0      | 3         | 4.8        | 3         | 2.0        |
| Professional           | 33       | 37.1     | 12        | 19.4       | 45        | 29.8       |
iterative process of (a) appraising the data, (b) creating codes (i.e. the most basic segment of the data that may be analysed meaningfully), and (c) recognising and reviewing broader themes (i.e. grouping the codes into semantically explicit categories). Thus, the outcome of this method is reporting of the concepts (i.e. codes) and patterns (i.e. themes) that emerged within the data (Braun & Clarke, 2006). The analysis process involved reiterative looping between the collected raw data and the

Table 1 (continued)

|                              | Phase 1 (N=89) |        | Phase 2 (N=62) |        | Total (N=151) |        |
|------------------------------|----------------|--------|----------------|--------|---------------|--------|
|                              | n  | %     | n  | %     | n  | %     |
| Clerical or administrative worker | 34 | 38.2  | 6  | 9.7   | 40 | 26.5  |
| Sales worker                 | 0  | 0.0   | 1  | 1.6   | 1  | 0.7   |
| Labourer                     | 0  | 0.0   | 1  | 1.6   | 1  | 0.7   |
| Machinery operator or driver | 0  | 0.0   | 2  | 3.2   | 2  | 1.3   |
| Technician or trade worker   | 0  | 0.0   | 2  | 3.2   | 2  | 1.3   |
| Other (not listed above)     | 0  | 0.0   | 4  | 6.5   | 4  | 2.6   |
| Non-response                  | 22 | 24.7  | 31 | 50    | 53 | 35.1  |
| Parental status              |    |        |    |        |    |        |
| No child                     | 40 | 44.9  | 0  | 0.0   | 40 | 26.5  |
| Younger children only        | 6  | 6.7   | 2  | 3.2   | 8  | 5.3   |
| School-aged children         | 15 | 16.9  | 16 | 25.8  | 31 | 20.5  |
| Adult children only          | 5  | 5.6   | 20 | 32.3  | 25 | 16.6  |
| Non-response                  | 23 | 25.8  | 24 | 38.7  | 47 | 31.1  |
| Region                       |    |        |    |        |    |        |
| Metropolitan                 | 89 | 100   | 37 | 59.7  | 126| 0.83  |
| Regional                     | 0  | 0.0   | 10 | 16.1  | 10 | 0.7   |
| Rural                        | 0  | 0.0   | 15 | 24.2  | 15 | 1.0   |
| Non-response                  | 0  | 0.0   | 0  | 0.0   | 0  | 0.0   |
| Taxpayer                     |    |        |    |        |    |        |
| No                           | 6  | 6.7   | 14 | 22.6  | 20 | 13.2  |
| Yes                          | 62 | 69.7  | 47 | 75.8  | 109| 72.2  |
| Non-response                  | 21 | 23.6  | 47 | 75.8  | 109| 72.2  |
| Resident status              |    |        |    |        |    |        |
| Citizen                      | 44 | 49.4  | 58 | 93.5  | 102| 67.5  |
| Permanent                    | 10 | 11.2  | 3  | 4.8   | 13 | 8.6   |
| Temporary                    | 13 | 14.6  | 1  | 1.6   | 14 | 9.3   |
| Non-response                  | 22 | 24.7  | 0  | 0.0   | 22 | 14.6  |

†Average fortnightly income after taxes and other deductions; < $1000 = Less than $1000; > $4000 = More than $4000; For the Phase 1 respondents, occupation data were reconstructed based on their education and income level, and all of them were categorised as ‘Metropolitan’ as they lived or worked in the metropolitan area. These two questions, occupation and region, were not included in Phase 1 survey; Response categories for education and occupation are those recognised by the Australian Bureau of Statistics’ (ABC) Census QuickStats (https://quickstats.censusdata.abs.gov.au/census_services/) as the most relevant and common categories reflecting the Australia society
developed codes and themes in order to reveal patterns that emerge directly from the participant quotes, rather than relying on latent tenuous interpretations (Boyatzis, 1998). Each of the codes represents a distinctive meaning but they may not be entirely mutually exclusive from the other codes because the smallest unit of interpretation (codes) may converge into the associated themes. The exactitude to detail in the participant quotes was enacted to retain consistency throughout the analysis. The direct quotes presented in the results section are a selected set, rather than all-inclusive, that is deemed to represent the central meaning of the identified codes or themes.

The first and third authors analysed the data separately, followed by four rounds of discussions and negotiations until we reached a complete agreement about the descriptors of the codes and emerging themes. Once the identified codes and themes were agreed upon among the three authors of this study, the first author read the entirety of the data as the final check if the emerged codes and themes are representative of the data as a whole. Finally, the codes and themes were evaluated against the initial theoretical framework adopted to guide the current study, which is presented in Section 5.

4 Results

Our data suggests a total of 11 codes, each of which is described in the following sections (the codes are presented in subheading).

4.1 Useful information to support learning

The respondents recognised NAPLAN’s value for its unique and in-depth information provided to students, parents, schools, and communities. Specifically, NAPLAN was viewed as a useful tool to identify areas requiring more learning support, compare individual students’ learning outcomes against the average, and monitor the performance gap between Indigenous and non-Indigenous students. NAPLAN data are seen as more specific in addressing learning issues of students than typical school reports can provide.

“It provides information that can be used as a guide or a checkpoint, useful for the students/parents, schools, communities” [Participant #17]

“I like it - it is the only way you can gauge your child against the average, and it very clearly shows what needs to be improved. Unlike the school reports which do not let you know what the problem is or how specifically to address any issues. Not even the teachers give out that in depth information. The NAPLAN is great” [Participant #27]

“NAPLAN has provided data to help us quantify the gap between Indigenous and non-Indigenous students’ literacy and numeracy” [Participant #64]

“It is important to identify the literacy and numeracy skills of young learners. There can be benefits to this testing to assist in identifying areas and students who may require more support in the learning process” [Participant #11]
4.2 National barometer of student and school performance

Several respondents in both Phases 1 and 2 found a standardised testing system necessary to nationally assess student and school performance. There was a broad recognition that NAPLAN allows students and schools to see their performance against the national curriculum and standards set by the Australian government and to guide themselves to identify strength and weakness.

“I think it is a good idea... we need a method of assessing how students are going in Australian schools” [Participant #24]

“It is necessary to track how the country’s children are performing” [Participant #13]

“We need a national barometer, that should be used as a guide for weakness and strength” [Participant #88]

“I think NAPLAN is a good tool to access the current level of student understanding in the [national] curriculum. It allows for schools to see if there are certain things to work and improve on the [national] standards of education” [Participant #29]

4.3 Assessment allowing school performance comparison

One of the unique functions of NAPLAN was recognised as providing comparative information by which school performance can be compared within the state and across the country. It was indicated that such an approach can be used to meet the school performance target and improve school performance and the educational system.

“In my opinion NAPLAN is very useful to evaluate not only individuals’ performance but also more importantly the ranking of the school compared to other schools in the state and the country. This data can then be used by the management to identify key points that will help improve the performance of the students and hence the performance of the school” [Participant #62]

“Each school should look at the NAPLAN results and see where there are strengths and weaknesses and work to a school performance target, reviewed annually” [Participant #111]

“It might be a good approach for a comparative assessment of the educational system across the institutions” [Participant #21]

4.4 Information not useful or beneficial

Some respondents in both Phases 1 and 2 expressed discontent because NAPLAN was seen as not generating relevant and timely information for stakeholders. NAPLAN was viewed as not providing additional information above and beyond what is already available in regular school-based assessments. Part of its failings was the delayed reporting time which prevents the information from being valuable to student learning.
“The test is basically used to measure the literacy and numeracy skills of young Australians, but schools already do pre and post-tests whose results go on end of year reports. I fail to see how testing students in this manner improves the quality of their education” [Participant #57]

“I see no evidence that standardised testing is raising educational standards across Australia (the opposite, in fact, if national comparisons are to be trusted)” [Participant #56]

“They [schools] should get results earlier to ensure they can use the info to help students to understand the areas they are unsure of. NAPLAN should be used to help students not schools” [Participant #74]

### 4.5 Individuals' strengths and preferences being ignored

The respondents expressed the concern that standardised testing such as NAPLAN ignores the unique strengths and learning style of individual students. They noted that not all students excel in academic subjects and there are different ways to learn. Standardised testing is unfair to those whose strengths were not in the mainstream subjects, and they may feel incompetent or behind.

“It doesn’t allow for the unique differences and strengths in children to be identified and encouraged” [Participant #08]

“I feel there are some students/schools who do not achieve best in this way” [Participant #32]

“I feel that in a way, the standardized testing makes it equal for everyone, since they are all taking the same questions. However, in a way, this also makes it unfair for the people whose niche does not lie in academics, but perhaps in something more abstract like music. These students who end up scoring more poorly than their peers may have the impression that they are stupid or incapable, when in fact, it’s just that they excel somewhere else” [Participant #30]

### 4.6 Teaching to the test

Many respondents mentioned NAPLAN as negatively incentivising the public and stakeholders by creating a counterproductive educational culture where schools prepare their students for this test, which potentially limits broader aspects of real learning. It appears that the respondents do not value rote learning to pass the test and doubted if actual learning can be encouraged for students when school and teachers teach to the test.

“A school should be teaching literacy and numeracy to all students aimed at their needs not the NAPLAN exam” [Participant #126]

“It’s terrible. Most schools teach to the test, say it doesn’t matter until they are a top performing school and then it is all that matters” [Participant #14]
“Standardised testing is harmful. It promotes rote learning to ‘pass’ a test and takes away from actual learning” [Participant #22]
“It distracts from real study and real learning while they are taught how to take a test instead of how to learn” [Participant #63]
“The school shouldn’t practice the NAPLAN for a month before the test.” [Participant #76]
“My view of NAPLAN is that it is an ineffective program that can actively demotivate and discourage students from trying and achieving at school. Whilst some standardised testing is necessary to track student progress the current NAPLAN system is failing teachers and students and influencing learning for NAPLAN rather than learning for learning’s sake” [Participant #61]
“It is only a memory test(s)… It does not actually test the student.” [Participant #104]

4.7 Other important skills are being ignored

Some respondents noted that students are missing out on other important skills that are far more important for youth to prepare for their future life. Under the school’s focus on demonstrating improved NAPLAN results, students are missing out on the opportunities to improve critical thinking or employability skills. The recognition of the importance of soft skills, such as resilience, respect, and tolerance, is also diminished in the current school culture. As such, NAPLAN is seen as limiting such holistic development.

“I have no positive views about NAPLAN, … … a test that doesn’t measure anything except their ability to parrot information at a particular day and time, without taking into account their critical thinking skills or their performance when not in a test environment” [Participant #63]
“I’m not a fan… …. I think we should be teaching children how to develop soft skills (e.g., resilience), not how to memorise a bunch of facts/figures to regurgitate onto some paper in a very specific classroom condition. With the world of work as it stands, I think that ‘employability’ skills are far more important to focus on” [Participant #25]
“From a personal perspective, I’ve verified the inverse relation between NAPLAN scores from primary schools and respect, tolerance and academic quality” [Participant #12]

4.8 Biased nature of standardised testing

Several respondents have noted the bias inherent in the standardised testing due to its inability to account for existing differences in language ability, culture, special needs, or learning styles. Overall, NAPLAN is seen as not equitable and lacks attention to diverse needs of individual students, which is in conflict with Australia’s value in schooling that aims to accommodate the needs of all students. NAPLAN was further seen as exacerbating exclusion and segregation which has broader implications beyond schooling.
“It could be argued there is inherent bias for these style of assessments as they do not take into account the individuals preferred learning and assessment style” [Participant #11]

“NAPLAN does not seem to take into account the fact that students in schools don’t just comprise the average but also include the likes of special needs students” [Participant #28]

“Until there is a level playing field, it’s a waste of resources” [Participant #68]

“Standardized testing doesn’t tell you anything important about an individual student. It doesn’t really tell you what you want to know about the learning or achievement of a cohort of students either, because those things are mixed up with language ability, fluency in the dominant culture, and many other factors” [Participant #03]

“NAPLAN is not equitable - it supports exclusion and segregation in schooling and beyond” [Participant #59]

4.9 Too much attention to the assessment itself

Some respondents pointed out that too much attention has been paid to the assessment itself. They noted the NAPLAN debate (i.e. whether it is useful or not) is increasing while the focus should be more on how to rectify the issues inherent in the assessment so that schooling can be more for young children’s holistic development.

“Important to have national standards but it perhaps receives more attention and focus than it should in order to achieve all round development of children” [Participant #31]

“Too much emphasis has been placed on one external exam - more focus has to be put on annual local exams and proper feedback to students” [Participant #20]

“It is disappointing that the debate is increasing with time rather than identifying an alternative or rectifying the defect” [Participant #51]

4.10 Stress and suffering of children

Strong negative feeling was expressed by the respondents in both phases about NAPLAN’s role in exerting stress, anxiety, and suffering onto children and parents. It was acknowledged that students’ exposure to unnecessary stress during schooling would negatively impact learning for enjoyment, feeling positive at school, and building confidence about their future.

“NAPLAN is a disgrace. all I’ve witnessed is very stressed children and very anxious parents as a result. We should be building a system based on learning for enjoyment, the excellence will come as a by-product of that” [Participant #100]

“Set NAPLAN on fire and set it adrift. It is a goal in and of itself that puts general and helpful learning on the backburner making children suffer” [Participant #80]

“Nap plan has shown to stress students out more. School should support and focus on students’ strengths to feel positive at school” [Participant #102]
“It scares students and places a fear on them that unless they achieve the right mark, they won’t be able to achieve their dreams” [Participant #63]

4.11 Original purposes being lost

The respondents recognised that the original purpose of NAPLAN has been lost largely in terms of guiding and informing individual students about their progress and what the next steps of learning should be. It is largely due to the government’s attempt to tie the funding to NAPLAN results and the schools’ desire to meet the accountability requirements. Thus, it appears that the respondents recognised that as long as NAPLAN is used for accountability, there is less use of NAPLAN to support student learning. Thus, discontent was expressed about the schools’ malpractice of coaching their students due to the government putting pressure on schools for accountability and funding.

“Should be revisited in terms of its original purpose. Should be used to provide information for individual students on how they are progressing and what specific needs are needed to advance to the next stage/level. Lessen its utilisation for accountability; should be more for supporting student learning.” [Participant #06]

“Tying school funding to NAPLAN scores has the potential to weaken schools with disadvantaged students, and so does diverting significant class time to test prep instead of real learning.” [Participant #03]

“I disagree with the Naplan system as many schools coach their students to attain good results for their school.” [Participant #108]

5 Discussion

The present study aimed to provide some, admittedly tentative, understanding of how the public’s perceptions about a mandatory national testing program such as NAPLAN may emerge. Despite the intense scrutiny and public interest in NAPLAN, there has been a relative paucity of scholarly work that delved into exploring cognitive mechanisms underlying the formation of public perception regarding national testing programs such as NAPLAN. Our two sets of qualitative data identified 11 codes (i.e. the most basic segment of the data that may be analysed meaningfully). In our effort to group together a set of codes that share semantically related concepts, we found three broader themes (see Fig. 1), labelled (a) informational utility, (b) value-driven opinion, and (c) broad consequences.

5.1 Informational utility

The first theme, informational utility, consists of the four codes (see Fig. 1). This theme suggests that when the participants think about NAPLAN, they judge it on its informational value. Specifically, this theme represents the participants’ perception of NAPLAN as a national barometer that can support learning, identify performance gaps, and allow comparative assessment of school performance. Overall, the
utility of NAPLAN was noted because only national testing can provide information to stakeholders and members of the general public, above and beyond the typical functions of schools. Thus, the comments reflected empirical support for the *incentive effect* in that NAPLAN was viewed as providing (or not providing) benefits and self-interest of stakeholders as parents/carers or members of the society.

Although some respondents mentioned that NAPLAN offers little information especially because of the three-month delay from the testing (May) to the reporting time (August), this theme contained more positive sentiment than negative views that were often highlighted in previous studies (e.g. Harris et al., 2013; Klenowski & Wyatt-Smith, 2012). There can be a few possibilities for why the views represented in this theme are not predominantly negative. First of all, this finding is somewhat logical because of the positive connotation underlying this theme which arose from the comments recognising the broad utility value of the national testing. Further, it should not be seen as surprising in the context of past research that noted that people do recognise both positive and negative consequences of NAPLAN although the media often feature negative stories of NAPLAN and the ‘failing’ teaching profession (Shine, 2015). For instance, Ragusa and Bousfield (2017) reported that over a quarter of their qualitative data contained both positive and negative comments. The Canvass Strategic Opinion Research (2013) also noted that the teaching profession’s views on NAPLAN became more positive when asked about specific functions and perceived utility of NAPLAN, for example, with regard to guiding school curriculum and classroom pedagogy or allocation of the school budget. Another possibility is that this theme may reflect the recent shift towards the recognition of broader communities. In particular, NAPLAN is seen as responding to the public’s right to know and contributing to parents’ engagement with their children in conversations and making informed decisions about their children’s future (McGaw et al., 2020). Considering these possibilities, it is not surprising that positive sentiment was expressed in this theme. Further, what this theme represents is that the positive perceptions about NAPLAN were closely related to its informational value. Thus, it implies that members of the general public may support the national testing program as long as it is perceived to generate unique and in-depth information that cannot be obtained elsewhere.
5.2 Value-driven opinion

The second theme, *value-driven opinion*, consists of the four codes (see Fig. 1). This theme demonstrates the participants’ concerns about the national testing’s inability to cater to individual students’ strengths and learning style and its lack of attention to diversity in student backgrounds in language, culture, or special needs. They also noted that other soft and employability skills necessary to prepare youth for life have been neglected under the pressure to demonstrate literacy or numeracy skills in the NAPLAN testing. The participants also expressed deep concern about the increasing trend in Australian schools adopting the malpractice of teaching to the test due to the performance pressure. Overall, the views expressed in this theme demonstrate some evidence for the *interpretive effect* in that the participants seem to interpret the functions and impact of NAPLAN, not in isolation as a testing instrument, but in the context of the core social/educational values. As the 2008 Melbourne Declaration on Educational Goals for Young Australians (MCEETYA, 2008) indicates, supporting individual students with diverse needs and addressing equity issues are the most important principles of educational values in Australia. This theme represents the participants’ disappointment in relation to achieving these national values.

We also found the criticisms under this theme expressed by our participants reflect the type of negative narratives most often recognised in past research, such as teaching to the test, ‘dumbing down’ of learning, disadvantaging certain groups of students (e.g. Brady, 2013) and marginalising higher-order thinking (e.g. Harris et al., 2013). Thus, it appears that there has been some broad consensus about the negative impacts of NAPLAN across different studies. The policy feedback literature suggests that policy failures tend to yield consensus among members of the public while the public tends not to agree on positive aspects of the policy (Piersson, 1993). Our study shows some support to this claim in that negative criticism found in our study is similar to what has been claimed in the past. Social and educational policy scholars (e.g. Jacobsen et al., 2014) have warned that negative public perception towards a particular program should be seen as a failure of that program when such negative views are stable over time. Given the almost identical nature of the criticisms surrounding NAPLAN espoused over the past decade the authorities responsible for the implementation of NAPLAN should be compelled to take these concerns seriously.

Overall, this theme reflects negative sentiments, frustration and discontent surrounding NAPLAN, and it is not surprising that positive aspects of NAPLAN did not emerge with respect to this theme. As mentioned above, the specific content of the criticism surrounding NAPLAN is not new. However, our study finding suggests an important implication that perhaps the root of specific criticisms about NAPLAN may be due to the public’s perception of the test’s misalignment with the dominant or ‘common-sense’ Australian values about what schooling should be (i.e. no harm to students, addressing individual students’ needs, caring for vulnerable populations). Thus, the present study findings call for the need to make concerted efforts to clearly articulate the core educational values held by the majority of the member of the public. It would also be essential to take into account the changing demographics of the society so that there should be a balance in the value articulation reflecting
both traditional Anglo-Western values (emphasizing individuals’ rights) and the
diverse social and educational values held by new members of the society.

5.3 Broad consequences

The third and last theme, consisting of the three codes (see Fig. 1), addresses broad consequences. The participants’ comments reflect the perception that NAPLAN has created a social-educational environment that causes stress and anxiety among children and parents. They also recognise too much attention given to the testing itself while losing sight of its main purpose of schooling in Australian society. Thus, the participants’ criticism about NAPLAN was not necessarily about the existence of the national testing itself but more to do with how it is (mis)used by stakeholders (e.g. schools) and government sectors. In particular, the federal government’s decision to make the school results available in the public domain, i.e. on the website, continues to dominate as the main concern in this study and as noted in many previous studies. It has been widely recognised that broad discontent with NAPLAN stems primarily from school comparison data being available in the public domain (e.g. Cumming et al., 2018; Joseph, 2018). It was suggested that while the assessment itself may not be critically deficient as a testing instrument (e.g. Lee et al., 2019), the way the government currently uses, disseminates, and shares the test results with the public remains a major issue (Wilson et al., 2021). As such, the finding that the negative consequences of NAPLAN were viewed with respect to its impact on the educational and societal climate gives rise to the empirical support for the institutional effect. Due to NAPLAN’s high visibility and traceability, the public will continue to call for its reform as long as the program’s main function appears to be derailed from its original intention because of the misuse of test information by the governing institution and schools.

In both cases of NCLB in the USA and NAPLAN in Australia, the school performance data have been made publicly available to members of the general public. Part of this institutional approach is to create “public pressure through publicising performance data” (Jacobsen et al., 2013, p. 360) with the hope that members of the public become aware of the system performance and perhaps put pressure on schools to improve (McDonnell, 2000; Jacobsen et al., 2013). Indeed, this approach has greatly contributed to the public debates about national testing program (often through media; Shine, 2015; Shine & Rogers, 2021) and will continue to play a significant role in the public’s perceptions about both positive and negative consequences of national testing.

5.4 Practical implications

In short, our study found that the opinions held by members of the public about NAPLAN are influenced not only by the perceived benefits to themselves (incentive effect) but also by their views on how the system-level entities such as schools and government interfere with the main function of schooling (institutional effect) and whether students and communities are treated fairly and served appropriately by the program (interpretive
Each of these elements was the impetus to generate the three research questions of this study: Would self-incentives explain perceptions about NAPLAN held by members of the general public (in relation to the incentive effect)? Would people’s perception of NAPLAN reflect the core educational values of Australia that emphasise the needs of individual students, social solidarities, and equity (in relation to the interpretive effect)? Would the program’s high visibility created by the institution contribute to the public’s interest and perception of NAPLAN (in relation to the institutional effect)? Our findings summarised in the codes and themes suggest that all these three questions were valid questions. Specifically, (a) the incentive effect was demonstrated in the public’s perceptions about the functional values of the instrument itself; (b) the interpretive effect was seen by the public’s comments about the meanings of the goals that the instrument is set out to achieve; and (c) the institutional effect was largely shown in the public’s disappointment in the other extraneous settings of NAPLAN that was imposed by the authorities (e.g. My School website, accountability discussions).

It should also be noted that these three effects are not independent of but intertwined with each other because individuals who perceive concrete benefits from a policy to themselves (incentive effect) would support the institution to maintain it and perhaps even mobilise their voices to be heard in the public domain, in the hope that the institution would take notice of their support and sustain the policy (institutional effect). Further, people’s normative expectations (interpretive effect) and understanding of the associated benefits (incentive effect) are likely to influence each other and further influence how fair (or unfair) they view the institutional use of the test information is in the context of informing the public about the purpose of the national testing (institutional effect).

The policy feedback literature identifies settings, instruments, and goals as three orders of policy change (Pierson, 1993). The first-order change requests the modification of the settings of the policy instruments. The second-order change alters the instruments as well as settings while keeping the policy goals intact. The third-order change requires a clear break from the past instruments, shifts in the settings, and re-orienting of the goals (Pierson, 1993, p. 614). It appears that in the case of NAPLAN, the public’s criticism was largely directed towards its settings and goals and less so towards the instrument itself. If the instrument itself was not highly criticised as this study demonstrated, the shift in the settings and re-defining the goals may be the necessary steps to address the widely recognised discontentment towards the national testing program.

In terms of clarifying the benefits of NAPLAN as a tool to guide and monitor student development, it may be useful for the institutions to clarify its function that goes beyond classroom-based assessments. Such functions include (a) providing the longitudinal link of student performance across primary and secondary schools (including transitional period between primary and secondary schools), (b) evaluating the performance outcomes using the cross-year continuous bands that allow self-checking and monitoring beyond the current year-level (e.g. a Year 5 student can see their performance reaching that of Year 9); and (c) reflecting student/school performance in lieu of the national and state/territory averages (which is not available in regular school reports).

In reviewing of the settings and goals of NAPLAN, there must be close collaboration between the institutions (government and schools) and stakeholders in their efforts to articulate the testing’s main purposes. It would be worthwhile for the government and school leaders to evaluate whether the educational policy surrounding
NAPLAN meets the expectations of different subpopulations within the society (e.g. indigenous communities, and rural areas). Some of the emerging social norms and values held by different ethic cultures and immigrant diaspora may also be explored for the program to have long-term sustainability.

5.5 Limitations

We note some limitations of the current study. First of all, we acknowledge that our sample is not representative of the NSW population. Therefore, the findings expressed in this study should be viewed as those of our participants who were willing to provide their opinion about NAPLAN. A more rigorous representative sampling procedure can be adopted in future and more comprehensive studies to ensure generalisable findings to broader communities. Also note that our sample is limited to those living in one state of Australia, NSW, and not across all states and territories of Australia. Second, it is possible that selection bias may have been present in our sample because the research participation was completely voluntary. However, the selection process may be taken as a necessary step because it indicates the participants’ willingness and comfort to provide the written responses to the topic being asked. Third, our study did not include the views of the most important group, i.e. students and teachers themselves. Fourth, we took a qualitative approach, which prevents us from quantification of the results (e.g. which effect may be more dominant in forming public opinion). Finally, the present study did not delve into potential subgroup differences. The policy feedback literature (e.g. Busemeyer, 2012; Pierson, 1993) suggests that self-interest tends to be influenced and shared by those who possess a similar skill portfolio, occupation, or educational attainment level. Thus, future studies may examine potential differences in perceiving the benefits of NAPLAN by different sectors of subpopulations within the Australian society.

6 Concluding remarks

The present study aimed to provide some tentative explanation for cognitive mechanisms of formation of the public perceptions surrounding a seemingly unpopular educational policy. When a national policy is perceived as a failure, “the complexity and multiplicity of policy interventions, combined with the uncertainty of the links between interventions and outcomes, will generally leave considerable room for dispute” (Pierson, 1993, p. 615). Thus, it would be critical to disentangle the root sources of the criticisms, which should be used for the development or refinement of the national testing program. Some of our recommendations include (a) providing a diverse range of incentives to promote self-interest; (b) shaping public narratives with the social norms and attitudes that are believed to be prevalent in society; and (c) managing the policy feedback loop between members of the public and institutions. It appears that the public support for NAPLAN may continue to suffer if its main purpose is poorly communicated to the public, its national testing benefits are broadly misunderstood, and the institution’s use of the test information is seen as unfair and unjustifiable.
## Appendix

Please see Table 2

### Table 2 Details of the sociodemographic background of the study participants in phases 1 and 2

| ID | Phase | Age | Gender | Education | Income     | Employment   | Occupation     | Parental status | Region          | Tax payer | Resident status |
|----|-------|-----|--------|-----------|------------|-------------|---------------|----------------|-----------------|------------|-----------------|
| 1  | 1     | 59  | F      | PhD       | $3000–$3999| Full-time   | Professional  | No child        | Metropolitan    | Yes        | Citizen         |
| 2  | 1     | 33  | M      | PhD       | $2500–$2999| Full-time   | Professional  | No child        | Metropolitan    | Yes        | Permanent       |
| 3  | 1     | 41  | F      | PhD       | $3000–$3999| Full-time   | Professional  | School-aged children | Metropolitan | Yes        | Citizen         |
| 4  | 1     | 45  | F      | PhD       | $3000–$3999| Full-time   | Professional  | No child        | Metropolitan    | Yes        | Temporary       |
| 5  | 1     | 29  | F      | PhD       | $2500–$2999| Full-time   | Professional  | No child        | Metropolitan    | Yes        | Temporary       |
| 6  | 1     | 40  | M      | PhD       | $3000–$3999| Full-time   | Professional  | School-aged children | Metropolitan | Yes        | Permanent       |
| 7  | 1     | 49  | F      | Bachelors | N/A        | Full-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 8  | 1     | 39  | F      | Bachelors | $1500–$1999| Full-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 9  | 1     | 41  | F      | PhD       | $3000–$3999| Full-time   | Professional  | Younger children only | Metropolitan | Yes        | Citizen         |
| 10 | 1     | 28  | M      | Bachelors | <$1000     | NILF        | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 11 | 1     | 32  | F      | Bachelors | $1500–$1999| Part-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 12 | 1     | 37  | M      | Masters   | $1500–$1999| Full-time   | Clerical      | School-aged children | Metropolitan | No         | Temporary       |
| 13 | 1     | 29  | M      | Masters   | $1000–$1499| Part-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 14 | 1     | 56  | F      | PhD       | $4000      | Full-time   | Professional  | School-aged children | Metropolitan | Yes        | Citizen         |
| 15 | 1     | 40  | F      | PhD       | $2000–$2499| Full-time   | Professional  | School-aged children | Metropolitan | Yes        | Citizen         |
| 16 | 1     | 22  | F      | Bachelors | <$1000     | Full-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 17 | 1     | 27  | F      | PhD       | $1000–$1499| Full-time   | Professional  | No child        | Metropolitan    | No         | Temporary       |
| 18 | 1     | 37  | M      | PhD       | $2000–$2499| Full-time   | Professional  | No child        | Metropolitan    | Yes        | Temporary       |
| 19 | 1     | 23  | F      | Masters   | $2000–$2499| Full-time   | Clerical      | No child        | Metropolitan    | Yes        | Citizen         |
| 20 | 1     | 58  | F      | Masters   | $2000–$2499| Full-time   | Clerical      | Adult children only | Metropolitan | Yes        | Citizen         |
| 21 | 1     | 42  | M      | Masters   | $1500–$1999| Full-time   | Clerical      | School-aged children | Metropolitan | No         | Temporary       |
| ID | Phase | Age | Gender | Education   | Income          | Employment | Occupation               | Parental status | Region         | Tax payer | Resident status |
|----|-------|-----|--------|-------------|-----------------|------------|--------------------------|-----------------|----------------|-----------|-----------------|
| 22 | 1     | 28  | F      | Bachelors   | $2000–$2499     | Part-time  | Clerical                 | No child        | Metropolitan   | Yes       | Citizen         |
| 23 | 1     | 35  | M      | PhD         | $3000–$3999     | Full-time  | Professional             | Younger children only | Metropolitan   | Yes       | Permanent       |
| 24 | 1     | 24  | F      | Bachelors   | < $1000         | Part-time  | Clerical                 | No child        | Metropolitan   | Yes       | Citizen         |
| 25 | 1     | 26  | M      | Bachelors   | < $1000         | Part-time  | Clerical                 | No child        | Metropolitan   | Yes       | Citizen         |
| 26 | 1     | 37  | M      | Masters     | $1500–$1999     | Part-time  | Clerical                 | No child        | Metropolitan   | Yes       | Citizen         |
| 27 | 1     | 47  | F      | PhD         | N/A             | Full-time  | Professional             | School-aged children | Metropolitan   | Yes       | Temporary       |
| 28 | 1     | 35  | F      | PhD         | $2500–$2999     | Full-time  | Professional             | No child        | Metropolitan   | Yes       | Temporary       |
| 29 | 1     | 28  | M      | Bachelors   | $1000–$1499     | Unemployed | Clerical                 | No child        | Metropolitan   | Yes       | Temporary       |
| 30 | 1     | 19  | F      | Year 10–12  | < $1000         | NILF       | Clerical                 | No child        | Metropolitan   | No        | Temporary       |
| 31 | 1     | 38  | F      | PhD         | $2000–$2499     | Part-time  | Professional             | Younger children only | Metropolitan   | Yes       | Citizen         |
| 32 | 1     | 35  | F      | PhD         | $2000–$2499     | Full-time  | Professional             | Younger children only | Metropolitan   | Yes       | Citizen         |
| 33 | 1     | 36  | F      | PhD         | $2000–$2499     | Full-time  | Professional             | Other            | Metropolitan   | Yes       | Citizen         |
| 34 | 1     | 37  | F      | PhD         | $3000–$3999     | Full-time  | Professional             | Younger children only | Metropolitan   | Yes       | Permanent       |
| 35 | 1     | 31  | F      | PhD         | $2500–$2999     | Full-time  | Professional             | No child         | Metropolitan   | Yes       | Citizen         |
| 36 | 1     | 19  | M      | Year 10–12  | < $1000         | NILF       | Clerical                 | No child        | Metropolitan   | No        | Temporary       |
| 37 | 1     | 27  | F      | Masters     | $1500–$1999     | Full-time  | Clerical                 | No child        | Metropolitan   | Yes       | Temporary       |
| 38 | 1     | N/A | F      | Bachelors   | N/A             | N/A        | Clerical                 | N/A             | Metropolitan   | Yes       | N/A             |
| 39 | 1     | 31  | F      | PhD         | $2500–$2999     | Full-time  | Professional             | No child        | Metropolitan   | Yes       | Citizen         |
| 40 | 1     | 31  | M      | Masters     | $2000–$2499     | Full-time  | Clerical                 | No child        | Metropolitan   | Yes       | Temporary       |
| 41 | 1     | 40  | M      | Bachelors   | $2000–$2499     | Full-time  | Clerical                 | School-aged children | Metropolitan   | Yes       | Permanent       |
| 42 | 1     | 54  | M      | Bachelors   | $2000–$2499     | Full-time  | Clerical                 | Adult children only | Metropolitan   | Yes       | Citizen         |
| 43 | 1     | 49  | M      | Masters     | $2000–$2499     | Full-time  | Clerical                 | School-aged children | Metropolitan   | Yes       | Citizen         |
| 44 | 1     | 23  | F      | Bachelors   | $1500–$1999     | Full-time  | Clerical                 | No child        | Metropolitan   | Yes       | Citizen         |
| 45 | 1     | 44  | M      | PhD         | $2500–$2999     | Full-time  | Professional             | No child        | Metropolitan   | Yes       | Citizen         |
| ID | Phase | Age | Gender | Education | Income      | Employment | Occupation              | Parental status | Region          | Tax payer | Resident status |
|----|-------|-----|--------|-----------|-------------|------------|-------------------------|----------------|----------------|-----------|-----------------|
| 46 | 1     | 29  | F      | Bachelors | $1500–$1999 | Part-time  | Clerical               | No child       | Metropolitan | Yes       | Citizen         |
| 47 | 1     | 39  | F      | PhD       | $3000–$3999 | Full-time  | Professional           | School-aged children | Metropolitan | Yes       | Citizen         |
| 48 | 1     | N/A | F      | PhD       | $3000–$3999 | Full-time  | Professional           | School-aged children | Metropolitan | Yes       | Citizen         |
| 49 | 1     | 30  | F      | Masters   | $2000–$2499 | Full-time  | Clerical               | No child       | Metropolitan | Yes       | Permanent       |
| 50 | 1     | 43  | M      | PhD       | <$1000     | Part-time  | Professional           | No child       | Metropolitan | Yes       | Citizen         |
| 51 | 1     | 47  | F      | PhD       | $3000–$3999 | Full-time  | Professional           | School-aged children | Metropolitan | Yes       | Permanent       |
| 52 | 1     | 31  | F      | Masters   | $1500–$1999 | Part-time  | Clerical               | No child       | Metropolitan | Yes       | Permanent       |
| 53 | 1     | N/A | F      | PhD       | $1500–$1999 | Full-time  | Professional           | No child       | Metropolitan | Yes       | Citizen         |
| 54 | 1     | 52  | F      | PhD       | $1500–$1999 | Full-time  | Professional           | Adult children only | Metropolitan | Yes       | Citizen         |
| 55 | 1     | 49  | M      | PhD       | $2500–$2999 | Full-time  | Professional           | No child       | Metropolitan | Yes       | Citizen         |
| 56 | 1     | 41  | F      | PhD       | $3000–$3999 | Full-time  | Professional           | School-aged children | Metropolitan | Yes       | Citizen         |
| 57 | 1     | 54  | F      | PhD       | >$4000     | Full-time  | Professional           | Adult children only | Metropolitan | Yes       | Citizen         |
| 58 | 1     | 34  | F      | PhD       | $2000–$2499 | Full-time  | Professional           | No child       | Metropolitan | Yes       | Permanent       |
| 59 | 1     | 42  | F      | Other     | >$4000     | Full-time  | Clerical               | No child       | Metropolitan | Yes       | Citizen         |
| 60 | 1     | 24  | F      | Bachelors | $1000–$1499 | Full-time  | Clerical               | No child       | Metropolitan | No        | Temporary       |
| 61 | 1     | 44  | F      | Masters   | $2000–$2499 | Part-time  | Clerical               | School-aged children | Metropolitan | Yes       | Citizen         |
| 62 | 1     | 33  | M      | PhD       | $1500–$1999 | Full-time  | Professional           | No child       | Metropolitan | Yes       | Citizen         |
| 63 | 1     | 28  | F      | Masters   | <$1000     | Full-time  | Clerical               | No child       | Metropolitan | Yes       | Citizen         |
| 64 | 1     | 39  | F      | Bachelors | <$1000     | Full-time  | Clerical               | No child       | Metropolitan | Yes       | Citizen         |
| 65 | 1     | 40  | F      | Masters   | >$4000     | Full-time  | Clerical               | School-aged children | Metropolitan | Yes       | Citizen         |
| 66 | 1     | 54  | F      | PhD       | >$4000     | Full-time  | Professional           | Adult children only | Metropolitan | Yes       | Citizen         |
| 67 | 1     | 23  | F      | Bachelors | $1000–$1499 | Part-time  | Clerical               | No child       | Metropolitan | Yes       | Permanent       |
| 68 | 2     | 67  | F      | Certificate | <$1000 | NILF | N/A | Adult children only | Metropolitan | No | Citizen |
| 69 | 2     | 22  | Other  | Diploma   | <$1000     | NILF | N/A | N/A | Metropolitan | Yes | Citizen |
| ID | Phase | Age  | Gender | Education     | Income      | Employment   | Occupation    | Parental status       | Region    | Tax payer | Resident status |
|----|-------|------|--------|---------------|-------------|--------------|--------------|----------------------|-----------|-----------|-----------------|
| 70 | 2     | 74   | M      | Certificate   | $2000–$2499 | NILF         | N/A          | Adult children only | Regional  | No        | Citizen          |
| 71 | 2     | 45   | M      | Masters       | $2500–$2999 | Full-time    | Professional | School-aged children | Metropolitan| Yes       | Citizen          |
| 72 | 2     | 45   | M      | Masters       | $3000–$3999 | Full-time    | Professional | School-aged children | Metropolitan| Yes       | Citizen          |
| 73 | 2     | 58   | F      | Year 10–12    | $2000–$2499 | NILF         | N/A          | Adult children only | Rural     | Yes       | Citizen          |
| 74 | 2     | 50   | F      | Vocational    | $1500–$1999 | NILF         | N/A          | School-aged children | Metropolitan| Yes       | Citizen          |
| 75 | 2     | 38   | F      | Vocational    | $1500–$1999 | Full-time    | Clerical     | School-aged children | Metropolitan| Yes       | Citizen          |
| 76 | 2     | 52   | F      | Year 10–12    | $1000–$1499 | NILF         | N/A          | School-aged children | Rural     | Yes       | Citizen          |
| 77 | 2     | 66   | M      | Bachelors     | $2000–$2499 | NILF         | N/A          | Adult children only | Regional  | Yes       | Citizen          |
| 78 | 2     | 71   | M      | Diploma       | $1000–$1499 | NILF         | N/A          | Adult children only | Rural     | No        | Citizen          |
| 79 | 2     | 41   | M      | Bachelors     | $2500–$2999 | Full-time    | Professional | N/A       | Metropolitan | Yes       | Citizen          |
| 80 | 2     | 42   | F      | Vocational    | <$1000      | Part-time    | Other        | N/A       | Metropolitan | Yes       | Citizen          |
| 81 | 2     | 40   | F      | Year 10–12    | <$1000      | Unemployed   | N/A          | N/A       | Rural       | Yes       | Citizen          |
| 82 | 2     | 43   | F      | Masters       | $2000–$2499 | Full-time    | Professional | School-aged children | Metropolitan| Yes       | Citizen          |
| 83 | 2     | 73   | M      | Vocational    | <$1000      | NILF         | N/A          | Adult children only | Rural     | No        | Citizen          |
| 84 | 2     | 52   | F      | Year 10–12    | <$1000      | NILF         | N/A          | School-aged children | Rural     | Yes       | Citizen          |
| 85 | 2     | 22   | F      | Bachelors     | $1000–$1499 | Part-time    | Sales worker | N/A       | Metropolitan | Yes       | Temporary         |
| 86 | 2     | 68   | F      | Certificate   | >$4000      | Part-time    | Clerical     | Adult children only | Rural     | Yes       | Citizen          |
| 87 | 2     | 56   | F      | Diploma       | $1000–$1499 | Part-time    | Other        | Adult children only | Metropolitan| Yes       | Citizen          |
| 88 | 2     | 64   | F      | Diploma       | $1500–$1999 | NILF         | N/A          | Adult children only | Regional  | No        | Citizen          |
| 89 | 2     | 33   | F      | Masters       | $1000–$1499 | Part-time    | Professional | N/A       | Metropolitan | Yes       | Citizen          |
| 90 | 2     | 49   | F      | Bachelors     | $1500–$1999 | Full-time    | Professional | N/A       | Metropolitan | Yes       | Citizen          |
| 91 | 2     | 52   | F      | Vocational    | $1500–$1999 | Full-time    | Clerical     | N/A       | Metropolitan | Yes       | Permanent         |
| 92 | 2     | 42   | F      | Year 10–12    | $1500–$1999 | Full-time    | Clerical     | School-aged children | Metropolitan| Yes       | Citizen          |
| 93 | 2     | 44   | M      | Bachelors     | $1000–$1499 | Full-time    | Technician   | School-aged children | Rural     | Yes       | Citizen          |
Table 2 (continued)

| ID | Phase | Age | Gender | Education | Income | Employment | Occupation | Parental status | Region          | Tax payer | Resident status |
|----|-------|-----|--------|-----------|--------|------------|------------|----------------|----------------|-----------|----------------|
| 94 | 2     | 74  | F      | Diploma   | < $1000| NILF       | N/A        | Adult children only | Metropolitan   | N/A      | Citizen         |
| 95 | 2     | 58  | M      | Vocational| $1000–$1499| Part-time | Machinery | N/A           | Regional       | Yes      | Permanent       |
| 96 | 2     | 51  | F      | Certificate| < $1000| Part-time | Clerical  | School-aged children | Metropolitan   | Yes      | Citizen         |
| 97 | 2     | 18  | M      | Year 10–12| $1000–$1499| Full-time | Machinery | N/A           | Metropolitan   | Yes      | Citizen         |
| 98 | 2     | 57  | M      | Vocational| $1000–$1499| Part-time | Technician | N/A           | Metropolitan   | Yes      | Citizen         |
| 99 | 2     | 25  | F      | Bachelors | $1500–$1999| Full-time | Other     | N/A           | Rural          | Yes      | Citizen         |
| 100| 2     | 31  | F      | Certificate| $2000–$2499| Part-time | Professional | Younger children only | Metropolitan   | Yes      | Citizen         |
| 101| 2     | 34  | F      | Bachelors | $1500–$1999| Part-time | Professional | Younger children only | Metropolitan   | Yes      | Citizen         |
| 102| 2     | 35  | F      | Bachelors | $1000–$1499| Full-time | Professional | School-aged children | Metropolitan   | Yes      | Citizen         |
| 103| 2     | 43  | M      | Vocational| < $1000| NILF       | N/A        | N/A           | Regional       | No       | Citizen         |
| 104| 2     | 66  | M      | Diploma   | $1500–$1999| NILF       | N/A        | N/A           | Rural          | No       | Citizen         |
| 105| 2     | 50  | F      | Year 10–12| $1500–$1999| Full-time | Labourer  | N/A           | Metropolitan   | Yes      | Citizen         |
| 106| 2     | 18  | M      | Year 10–12| < $1000| NILF       | N/A        | N/A           | Rural          | No       | Citizen         |
| 107| 2     | 20  | F      | Year 10–12| $1000–$1499| Full-time | Manager   | N/A           | Rural          | Yes      | Citizen         |
| 108| 2     | 66  | F      | Diploma   | $1000–$1499| NILF       | N/A        | Adult children only | Rural          | Yes      | Citizen         |
| 109| 2     | 28  | F      | Year 10–12| < $1000| Unemployed | N/A        | School-aged children | Regional       | No       | Citizen         |
| 110| 2     | 39  | F      | Vocational| $1500–$1999| NILF       | N/A        | School-aged children | Rural          | No       | Citizen         |
| 111| 2     | 58  | M      | Bachelors | > $4000| NILF       | N/A        | N/A           | Metropolitan   | Yes      | Citizen         |
| 112| 2     | 73  | M      | Diploma   | $2500–$2999| NILF       | N/A        | Adult children only | Metropolitan   | Yes      | Citizen         |
| 113| 2     | 79  | M      | Diploma   | $3000–$3999| NILF       | N/A        | Adult children only | Metropolitan   | Yes      | Citizen         |
| 114| 2     | 64  | M      | Masters   | $1500–$1999| Part-time | Manager   | Adult children only | Metropolitan   | Yes      | Citizen         |
| 115| 2     | 53  | M      | Bachelors | $3000–$3999| Full-time | Manager   | School-aged children | Metropolitan   | Yes      | Citizen         |
| 116| 2     | 69  | F      | Vocational| $1000–$1499| NILF       | N/A        | Adult children only | Regional       | No       | Citizen         |
| 117| 2     | 55  | F      | Masters   | $1000–$1499| NILF       | N/A        | N/A           | Regional       | Yes      | Citizen         |
| ID  | Phase | Age | Gender | Education        | Income        | Employment  | Occupation         | Parental status         | Region | Tax payer | Resident status |
|-----|-------|-----|--------|------------------|---------------|-------------|-------------------|-------------------------|--------|-----------|-----------------|
| 118 | 2     | 39  | M      | Diploma          | $1500–$1999   | Part-time   | Other             | Adult children only       | Metropolitan | No        | Citizen           |
| 119 | 2     | 18  | F      | Year 10–12       | $3000–$3999   | Unemployed  | N/A               | N/A                      | Metropolitan | Yes       | Citizen           |
| 120 | 2     | 63  | M      | Masters          | > $4000       | NILF        | N/A               | Adult children only       | Metropolitan | Yes       | Citizen           |
| 121 | 2     | 60  | F      | Vocational       | < $1000       | NILF        | N/A               | Adult children only       | Metropolitan | No        | Citizen           |
| 122 | 2     | 47  | M      | Year 10–12       | $1000–$1499   | NILF        | N/A               | N/A                      | Metropolitan | Yes       | Citizen           |
| 123 | 2     | 47  | M      | Year 10–12       | $1000–$1499   | NILF        | N/A               | N/A                      | Metropolitan | Yes       | Citizen           |
| 124 | 2     | 47  | F      | Bachelors        | $1500–$1999   | Full-time   | Clerical          | School-aged children      | Metropolitan | Yes       | Citizen           |
| 125 | 2     | 43  | M      | Bachelors        | $2000–$2499   | Part-time   | Professional      | N/A                      | Metropolitan | Yes       | Citizen           |
| 126 | 2     | 40  | F      | Certificate      | $1500–$1999   | Part-time   | Professional      | School-aged children      | Regional   | Yes       | Citizen           |
| 127 | 2     | 66  | M      | Year 10–12       | $1500–$1999   | NILF        | N/A               | Adult children only       | Regional   | Yes       | Citizen           |
| 128 | 2     | 30  | F      | Bachelors        | $1500–$1999   | Full-time   | Professional      | N/A                      | Metropolitan | Yes       | Citizen           |
| 129 | 2     | 68  | F      | Other            | $1000–$1499   | NILF        | N/A               | Adult children only       | Rural      | No        | Permanent         |

All categories for each of the demographic variables are listed in Table 1: F, Female; M, Male; < $1000 = Less than $1000; > $4000 = More than $4,000; Vocational, Vocational Certificate; Certificate, Graduate Diploma or Certificate; Income is the average fortnightly income after taxes and other deductions; In the Phase 1 data, 22 participants did not provide more than half of the demographic information variables and are not listed in this table; For the Phase 1 respondents, occupation data were reconstructed based on their education and income level, and all of them were categorised as ‘Metropolitan’ as they lived or worked in the metropolitan area as these two questions were not included in Phase 1 survey; N/A, Missing; NILF, ‘not in the labour force’
Funding  Open Access funding enabled and organized by CAUL and its Member Institutions This work was supported by the Gonski Institute for Education, University of New South Wales, Sydney Australia [Project Title: Beliefs about Educational Equity].

Data availability  The data that support the findings of this study are available from the corresponding author upon request.

Declarations

Conflict of interest  The authors declare no competing interests.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Berkman, M. B., & Plutzer, E. (2005). Ten thousand democracies. Georgetown University Press.
Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. Cambridge Journal of Education, 41(3), 287–302.
Boyatzis, R. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.
Brady, L. (2013). NAPLAN: Critiquing the criticisms. Curriculum and Teaching, 28(1), 47–55.
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.
Buckendahl, C. W. (2016). Public perceptions about assessment in education. In G. T. L. Brown & L. R. Harris (Eds.), Handbook of human and social conditions in assessment (pp. 454–471). Routledge.
Busemeyer, M. R. (2012). Inequality and the political economy of education: An analysis of individual preferences in OECD countries. Journal of European Social Policy, 22(3), 219–240.
Busemeyer, M. R., & Garritzmann, J. L. (2017). Academic, vocational or general? An analysis of public opinion towards education policies with evidence from a new comparative survey. Journal of European Social Policy, 27(4), 373–386.
Canvass Strategic Opinion Research. (2013). Primary principals: Perspectives on NAPLAN testing and assessment. Authors.
Cumming, J., Wyatt-Smith, C., & Colbert, P. (2016). Students ‘at risk’ and the National Assessment Program Literacy and Numeracy (NAPLAN): The ‘collateral damage’. In R. Lingard, G. Thompson, & S. Sellar (Eds.), National testing and its effects: Evidence from Australia (pp. 126–138). Oxon, England: Routledge.
Cumming, J., Jackson, C., Day, C., Maxwell, G., Adie, L., Lingard, B., ... & Heck, E. (2018). 2018 Queensland NAPLAN review: School and system perceptions report and literature review. Institute for Learning Sciences and Teacher Education, Australian Catholic University and the Queensland Government. https://qed.qld.gov.au/programinitiatives/education/Documents/naplan-2018-school-perceptions-report.pdf
Elliott, J., Stankov, L., Lee, J., & Beckmann, J. F. (2019). What did PISA and TIMSS ever do for us?: The potential of large scale datasets for understanding and improving educational practice. Comparative Education, 55(1), 133–155.
Fischman, G. E., Topper, A. M., Silova, L., Goebel, J., & Holloway, J. L. (2019). Examining the influence of international large-scale assessments on national education policies. *Journal of Education Policy, 34*(4), 470–499.

Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied Thematic Analysis*. Sage.

Hardy, I. (2014). A logic of appropriation: Enacting national testing (NAPLAN) in Australia. *Journal of Education Policy, 29*(1), 1–18.

Hargreaves, A. (2020). Large-scale assessments and their effects: The case of mid-stakes tests in Ontario. *Journal of Educational Change, 21*(3), 393–420.

Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly, 42*(1), 3–41.

Harris, P., Chinnappan, M., Castleton, G., Carter, J., De Courcy, M., & Barnett, J. (2013). Impact and consequence of Australia’s national assessment program-literacy and numeracy (NAPLAN)—using research evidence to inform improvement. *TESOL in Context, 23*(1/2), 30.

Heinström, J. (2010). *From fear to flow: Personality and information interaction*. Chandos.

Herman, J. L., & Golan, S. (1993). The effects of standardized testing on teaching and schools. *Educational Measurement: Issues and Practice, 12*(4), 20–25.

Jacobs, A. M., & Weaver, R. K. (2015). When policies undo themselves: Self-undermining feedback as a source of policy change. *Goverance, 28*(4), 441–457.

Jacobsen, R., Saultz, A., & Snyder, J. W. (2013). When accountability strategies collide: Do policy changes that raise accountability standards also erode public satisfaction? *Educational Policy, 27*(2), 360–389.

Jacobsen, R., Snyder, J. W., & Saultz, A. (2014). Informing or shaping public opinion? The influence of school accountability data format on public perceptions of school quality. *American Journal of Education, 121*(1), 1–27.

Joseph, B. (2018). *Why we need NAPLAN*. The Center for Independent Studies (CIS) Research Reports (RR36). Sydney, Australia. https://www.cis.org.au/wp-content/uploads/2018/05/rr36.pdf

Kearns, L. L. (2011). High-stakes standardized testing and marginalized youth: An examination of the impact on those who fail. *Canadian Journal of Education, 34*(2), 112–130.

Kelly, P., Andreasen, K. E., Kousholt, K., McNess, E., & Ydesen, C. (2018). Education governance and standardised tests in Denmark and England. *Journal of Education Policy, 33*(6), 739–758.

Klenowski, V., & Wyatt-Smith, C. (2012). The impact of high stakes testing: The Australian story. *Assessment in Education: Principles, Policy & Practice, 19*, 65–79.

Kohn, A. (2000). Standardized testing and its victims. *Education Week, 20*(4), 46–47.

Lee, J., McArthur, W., & Ellis, N. J. (2019). NAPLAN versus in-school assessment: How similar or different are students’ results? *Curriculum and Teaching, 34*(2), 5–25.

Levin, H. M. (1987). Education as a public and private good. *Journal of Policy Analysis and Management, 6*(4), 628–641.

Lewis, S., & Hardy, I. (2015). Funding, reputation and targets: The discursive logics of high-stakes testing. *Cambridge Journal of Education, 45*(2), 245–264.

Lingard, B., & Sellar, S. (2016). The changing organizational and global significance of the OECD’s education work. In K. Mundy, A. Green, B. Lingard, & T. Verger (Eds.), *International handbook of globalization and education policy* (pp. 357–373). Blackwells.

Louden, W. (2019). *NAPLAN Reporting Review*. Prepared for COAG Education Council. Education Services Australia on behalf of the Education Council. https://files.eric.ed.gov/fulltext/ED599927.pdf

MacKay, D. (1969). *Information, mechanism and meaning*. MIT Press.

Mattei, P. (2012). Raising educational standards: National testing of pupils in the United Kingdom, 1988–2009. *Policy Studies, 33*(3), 231–247.

McDonnell, L. M. (2000). Defining democratic purposes. In L. M. McDonnell, P. M. Timpane, & R. Benjamin (Eds.), *Rediscovering the democratic purposes of education* (pp. 1–20). University of Kansas Press.

MCEETYA. (1989). *The Hobart declaration on schooling*. Melbourne: The Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA).

MCEETYA (2008). *The Melbourne declaration on educational goals for young Australians*. Melbourne: The Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA).

McGaw, B., Louden, W., & Wyatt-Smith, C. (2020). *NAPLAN review final report*. State of New South Wales (Department of Education), State of Queensland (Department of Education), State of Victoria.
Mills, M., Monk, S., Keddie, A., Renshaw, P., Christie, P., Geelan, D., & Gowlett, C. (2014). Differentiated learning: From policy to classroom. *Oxford Review of Education, 40*(3), 331–348.

Musgrave, R. A. (1959). *The theory of public finance*. A study in public economy.

Perrin, A. J., & McFarland, K. (2011). Social theory and public opinion. *Annual Review of Sociology, 37*(1), 87–107.

Pierson, P. (1993). When effect becomes cause: Policy feedback and political change. *World Politics, 45*(4), 595–628.

Pizmony-Levy, O., & Bjorklund, P., Jr. (2018). International assessments of student achievement and public confidence in education: Evidence from a cross-national study. *Oxford Review of Education, 44*(2), 239–257.

Pizmony-Levy, O., Doan, L., Carmona, J., & Kessler, E. (2018). The public and international assessments. In B. Maddox (Ed.), *International large-scale assessments in education: Insider research perspectives* (pp. 197–218). Bloomsbury Publishing.

Ragusa, A. T., & Bousfield, K. (2017). ‘It’s not the test, it’s how it’s used!’ Critical analysis of public response to NAPLAN and My School Senate Inquiry. *British Journal of Sociology of Education, 38*(3), 265–286.

Ravitch, D. (2016). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.

Reid, A. (2020). *Here’s what happens when NAPLAN becomes more than an education tool*. Sydney Morning Herald. https://www.smh.com.au/education/heres-what-happens-when-naplan-becomes-more-than-an-education-tool-20190917-p52s3i.html

Rose, J., Low-Choy, S., Singh, P., & Vasco, D. (2020). NAPLAN discourses: A systematic review after the first decade. *Discourse Studies in the Cultural Politics of Education, 41*(6), 871–886.

Sellar, S., & Lingard, B. (2013). Looking East: Shanghai, PISA 2009 and the reconstitution of reference societies in the global education policy field. *Comparative Education, 49*(4), 464–485.

Shine, K. (2015). Are Australian teachers making the grade? A study of news coverage of NAPLAN testing. *Media International Australia, 154*(1), 25–33.

Shine, K., & Rogers, S. L. (2021). Parents and teachers’ perceptions of the tone and emotional impact of education news coverage. *Journalism and Media, 2*(2), 193–207.

Svallfors, S. (Ed.). (2012). *Contested welfare states: Welfare attitudes in Europe and beyond*. Stanford, CA: Stanford University Press.

Wilson, R., Piccoli, A., Hargreaves, A., Ng, P. T., & Sahlberg, P. (2021). *Putting students first: Moving on from NAPLAN to a new educational assessment system* (The Gonski Institute Policy Paper #2-2021). UNSW Gonski Institute.

Wyatt-Smith, C., & Klenowski, V. (2010). Role and purpose of standards in the context of national curriculum and assessment reform for accountability, improvement and equity in student learning. *Curriculum Perspectives, 30*(3), 37–47.

Zadkovich, G. (2017). NAPLAN’s failure prompts push to replace testing. *Education, 98*(5). https://www.proquest.com/openview/55a4fa11df42ddfdd7664e5b23accf771?pq-origsite=gsl&cbl=27966

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.