

The Iterative Survey Adaptation of The Imitation Game: A survey approach to defining group identity through analysis of distinctiveness

Methodological Innovations

1–12

© The Author(s) 2024

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20597991241285802

journals.sagepub.com/home/mioFay Cosgrove 

Abstract

This article reports on an original adaptation of the Imitation Game into survey form; the Iterative Survey. Like the original Game, the Iterative Survey can capture the ‘groupishness’ of social groups through analysis of distinctiveness but it can be implemented more affordably and conveniently than the Game making it more useful for researchers with limited resources. This paper outlines the origins of this novel method, and its implementation is illustrated using an exploration of maths anxiety in primary school teachers in Wales. The method privileges participant voices and its asynchronous nature may allow access to groups that would be difficult to recruit for a standard Imitation Game. The method’s rigour and ability to capture ‘groupishness’ are tentatively evaluated. The social scientific value of this novel survey adaptation is discussed with reference to wider social science. The article ends with a discussion of the need for further development of the method.

Keywords

Imitation Game, Iterative Survey, maths anxiety, teachers, primary, group identity, groupishness

Introduction

The Imitation Game (Collins et al., 2006) collects both qualitative and quantitative data and can be used as a mixed methods procedure across a wide range of disciplines. The pilot reported in this article introduces a new Iterative Survey adaptation of the Imitation Game that will be of value to social research. This project, the first to implement the Iterative Survey version of the Game, asks whether the adaptation retains the Game’s ability to capture the ‘groupishness’ (Evans et al., 2019) of social groups by identifying elements of group membership that are both shared within, and distinctive of, a social group. This paper reports the use of this novel adaptation in the context of exploring maths anxiety (MA) in primary school teachers in Wales¹. Published social science literature does not outline any attempt to convert a synchronous Imitation Game, nor any other kind of research game, into an asynchronous survey (see Appendix A) making this methodological approach novel.

The Imitation Game can capture ‘shared normative expectations – typically narrower in nature than the diversity displayed by individual group members’ (Evans et al., 2019: 1561). In other words, take a hypothetical group member and record all their experiences and feelings about group membership. Repeat for other members of the same group and compare the lists. Some elements will be idiosyncratic and some will be shared. The shared experiences encapsulate ‘groupishness’, and they ‘transcend the individual’ (Evans et al., 2019: 1561). Compared with the aggregated list, the list of shared experiences will be shorter and more representative of the group (because now the list only contains commonalities). Some entries

Cardiff University, Cardiff, UK

Corresponding author:

Fay Cosgrove, School of Social Sciences, Cardiff University, Glamorgan Building, Cardiff, CF10 3AT, UK.

Email: cosgrovefh@cardiff.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of

the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

on the list will also be common to unrelated groups while other entries will be distinctive of the target group.

The original Imitation Game is played synchronously by participants who are either members ('in-group') or non-members ('out-group') of the target group. The Game requires a bespoke website to facilitate anonymous, synchronous communication between players. The adapted survey version uses multiple iterations of a web survey, with each iteration informing the next. Reported here is the first use of the survey version; as a small-scale exploration of concepts of MA held by primary school teachers in Wales. The primary teacher population was deemed suitable for the first use of this mini-method due to having good literacy levels and widespread internet access (Welsh Government, 2022).

After describing the genesis of the method, the context of the current research will be summarised. Next, the adaptation from the original Imitation Game to Iterative Survey will be outlined. The article then focuses solely on the Iterative Survey, the method's ability to capture groupishness and its ethical implications. Finally, future developments of the survey adaptation will be considered.

The origins of the Imitation Game

The Imitation Game was derived from an early twentieth century parlour game that also inspired Alan Turing's test of artificial intelligence, known as the Turing Test² (Turing, 1950). The Imitation Game was first developed into a method for social science research by Collins et al. (2006). In a single Game, three players are needed; two in-group participants and one out-group participant. One of the in-group participants becomes the 'judge' while the other takes on the role of the 'non-pretender'. The judge asks a question designed to help them determine which player is a fellow in-group member and which player is merely pretending to be so (the out-group, 'pretender'). The judge must make the decision based on only the anonymised answers received from each of the players. Collins et al. (2006) discuss the implications in terms of types of expertise and the results of the Game can be used to assess how well the out-group understands the experience of the in-group and to identify shared yet distinctive elements of group membership (groupishness). It is this latter ability which is retained by the Iterative Survey.

This article reports on research in which the judge and non-pretender were both maths-anxious primary school teachers in Wales while the pretender was a primary teacher who was not maths-anxious (see Figure 1). The judge (a maths-anxious teacher) created a question they thought would help them determine the pretender (the out-group, non-maths anxious teacher) from the non-pretender (the in-group, maths-anxious teacher), knowing they would receive only anonymous answers to help them

decide which was which. In doing so, they were selecting an aspect of MA they thought other maths-anxious teachers would recognise yet which would not be shared by the out-group pretender, making it distinctive to those teachers who experience MA.

Previous uses of the Imitation Game

The Imitation Game has been used to explore a variety of social topics. Some games have been played between groups that co-exist but have no overt need to understand one another. For example, Segersven and Heino (2019) played the Game with people who had type 1 diabetes and those who did not. Collins et al. (2019) used the Game to research Scottish and English national identities. Evans et al. (2019) played it to look at performance of gender identity and Arminen et al. (2019) used it to compare the understandings of active Christians and non-Christians. Collins (2016) explored the expertise of gravitational wave physicists to non-gravitational wave physicists.

The Imitation Game has also been played with groups that have some kind of (often bi-directional) dependent relationship wherein one group is expected (but not always demonstrated) to have an understanding of the other. For example, Wehrens and Walters (2018) used the Game to evaluate the understandings of the lifeworlds of people with eating disorders and their therapists. The Game has also been used to examine the understandings of professional dieticians involved in the care of people living with Coeliac Disease (Boyce, 2009; Evans and Crocker, 2013). Segersven et al. (2024) used it to explore the hypothesised broader cultural competence of second generation Finnish Somalis compared with members of the native Finnish majority.

The relationship between the groups reported in this article (maths-anxious primary teachers and non-maths-anxious primary teachers) is probably best considered to be the first, co-existent type. While there are some non-maths-anxious teachers for whom it is desirable to understand their maths-anxious colleagues, namely those who have responsibility for supporting anxious teachers to develop their maths practice (head teachers, deputy heads, heads of department, etc.), this is not the case for most non-maths-anxious teachers. For the majority, they may work with colleagues who are anxious about maths without even realising that is the case if those maths-anxious teachers don't voluntarily disclose their feelings. Regardless of the exact relationship between the groups, the Imitation Game is able to identify shared yet distinctive elements of group membership and therefore there is reason to hope that the Iterative Survey shares this ability to identify 'groupishness' in social settings.

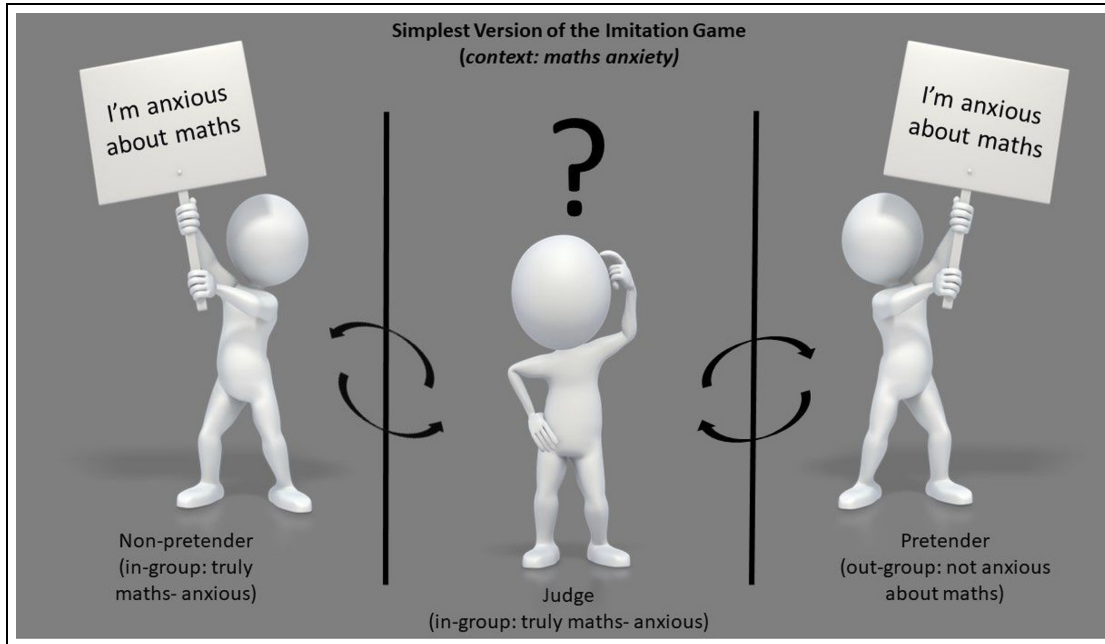


Figure 1. Simple version of an Imitation Game exploring maths anxiety.

Simultaneous gathering and evaluation of data

Data collected during a Game could be considered analogous to the information collected in solving the ‘*Liar Paradox*’ (Parsons, 1974). Two guards stand in front of a pair of doors, one leading to freedom and one to certain doom. One guard always tells the truth while the other always lies. The interrogator can ask only one question. The solution is to ask either guard which door the other would say leads to freedom and then take the alternative door (a lie about the truth and the truth about a lie both representing a falsehood). While neither the Imitation Game nor its survey adaptation include logical paradoxes, they do involve the simultaneous gathering and evaluation of data like that seen in the Liar Paradox: challenging a judge to pose a question elicits not just an element of groupishness, but one which the judge has selected for its potential assistance in discerning between the pretender and non-pretender, implicitly adding the second layer of information about perceived veracity and groupishness. Collins and Evans (2017) refer to group members revealing this layer of utility as, ‘persuading the thermometers to read themselves’ (Collins and Evans, 2017: 334). The survey adaptation, introduced in this paper, also enables ‘participants to make their [most distinctive] taken-for-granted assumptions and practices more visible’ (Collins and Evans, 2017: 335).

Substantive context

MA is more than a dislike of maths; it is a discrete form of anxiety often defined as,

feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in ordinary and academic situations. (Richardson and Suinn, 1972: 551)

The topic of more than 60 years of research (Dowker et al., 2016), MA has been considered part of the ‘international dilemma’ in maths and may be contributing to Western society’s acceptance of poor number skills (Chinn, 2020: 74) and disinterest in Science, Technology, Engineering and Maths careers. Often developed during primary school (Hill et al., 2016), MA is negatively correlated with maths attainment (Wang et al., 2015), especially for more able pupils (Foley et al., 2017) and, according to some, for girls more than boys (Hill et al., 2016; Organisation for Economic Co-operation and Development, 2015, etc.). Causes of MA are likely to be multiple, encompassing:

- individual factors (e.g. working memory capacity, Yucedag-Ozcan and Brewer, 2011),
- attributional style (Fernández-Sogorb et al., 2020),
- teacher and parent influence (e.g. Maloney et al., 2015; Schaeffer et al., 2020),
- the curriculum and school environment (e.g. Ashcraft et al., 2007; Buckley and Sullivan, 2023),
- and societal attitudes (Chinn, 2020).

MA is experienced by adults as well as children, in fact trainee primary teachers experience more MA than students studying degree programmes in non-education

disciplines (Artemenko et al., 2021; Hembree, 1990). Teacher MA has a negative impact on pupil maths learning (Beilock et al., 2010; Schaeffer et al., 2020), making understanding of teacher MA important for maximising children's numeracy.

As will be seen, the Iterative Survey version of the Game retains the Game's ability to capture distinctive, shared elements of MA while also being easily scalable and both quick and inexpensive to implement.

Adaptation from Game to Iterative Survey

The Imitation Game was selected for its ability to identify aspects of groupishness that offer the most promise of discernment between in- and out-group, in this case in order to inform future research on the measurement of MA in teachers. The Game was adapted into an online Iterative Survey in order to facilitate easy recruitment of primary school teachers from across Wales. The online nature of the survey avoided the need for the researcher to travel around the country and reduced scheduling concerns for participants due to its asynchronous nature.

The method used by Evans and Crocker (2013) provided an a priori reason to expect potential success for the Iterative Survey. They conducted an asynchronous version of the Game entirely by email. This is the only known example of a social science research game being conducted asynchronously. Judge-posed questions were emailed to the researchers, who anonymised them and emailed them to in-group and out-group players. Once answers were returned, they were assembled into appropriate sets and returned to the judges who made their identifications by email to the researchers. This method of communication proved time-consuming and somewhat onerous for the researchers (Evans, 2023). For this reason, and due to concerns about it being prone to researcher error, the survey approach was conceived and developed.

Item types

An iterative, branching survey was created in which there are three main item types plus a declaration (Q1) which separates in-group players from out-group players (see Figure 2). Those who declare themselves in-group (in this case by stating that they are maths-anxious), are presented with all three remaining item types. The first substantive item (Q2), which is thought to be the least cognitively demanding (Callegaro et al., 2015), is a multiple choice item that asks the participant to play one part of the judge's role: they are shown a question posed by a judge from a previous iteration of the survey and two corresponding answers (one answer from a pretender and one from a non-pretender). They are asked which they think came from the in-group, non-pretender (a truly maths-anxious teacher).

As can be seen from Figure 2, this item is necessarily wordy such that the participant understands the scenario. It is thought that the use of bold font and bullet points in this carefully formatted item helps to make it more accessible than a block of plain text (Miniukovich et al., 2017). While this item is lengthy, the current data suggests it is the item answered most quickly by participants and results in the least discontinuation, and so it should continue to appear first.

There follows an item type (Q3&4) which asks the in-group participant (who is maths-anxious) to play as a non-pretender and answer a question posed by a judge from a previous iteration. This item uses an open-ended text box. While open-ended text boxes in surveys can be more time-consuming to answer and be cognitively burdensome (Callegaro et al., 2015), this item was shown to be less time-consuming than the item that follows it. This is thought to be because participants need only consider their own experiences of MA, rather than both pretender's and non-pretender's potential answers.

The final item type (Q5) shown to in-group (maths-anxious) participants asks them to perform the other part of the judge's role and pose a new question which they think would help them distinguish between pretender and non-pretender, just as they would in the original Game. The instructions that accompany this question are shorter than those for the first item type, yet this item was found to be the most time-consuming for participants. Despite now being familiar with the scenario, this item is the most cognitively demanding, since it uses an open-ended text box and because there is a fairly complex scenario to consider: participants may be selecting from a possibly broad range of aspects of group membership. When this question was presented earlier in the anxious branch, it resulted in the most discontinuation.

In the original Game, the same person (playing as in-group judge) would pose a question and attempt to identify which player also belonged to the in-group based on the two answers to their questions. In the Iterative Survey, the judge who poses a question is not the same as the one who uses its answers to identify the players. By using two different judges in this way, the requirement for the authoring judge to remain available until after both players have contributed their answers is removed. This variation is vital to the adaptation of the Game into a survey and is the key to unlocking the asynchronicity offered by the Iterative Survey.

Those who declared themselves out-group (in this case stating that they are not anxious about maths) are given only one further item type (Q6&7). It is almost identical to the second item type on the anxious branch (Q3&4), but participants are asked to pretend that they *are* anxious about maths, thus fulfilling the role of the pretender. It is important that these questions match those shown on the anxious branch so that in the next iteration, the new

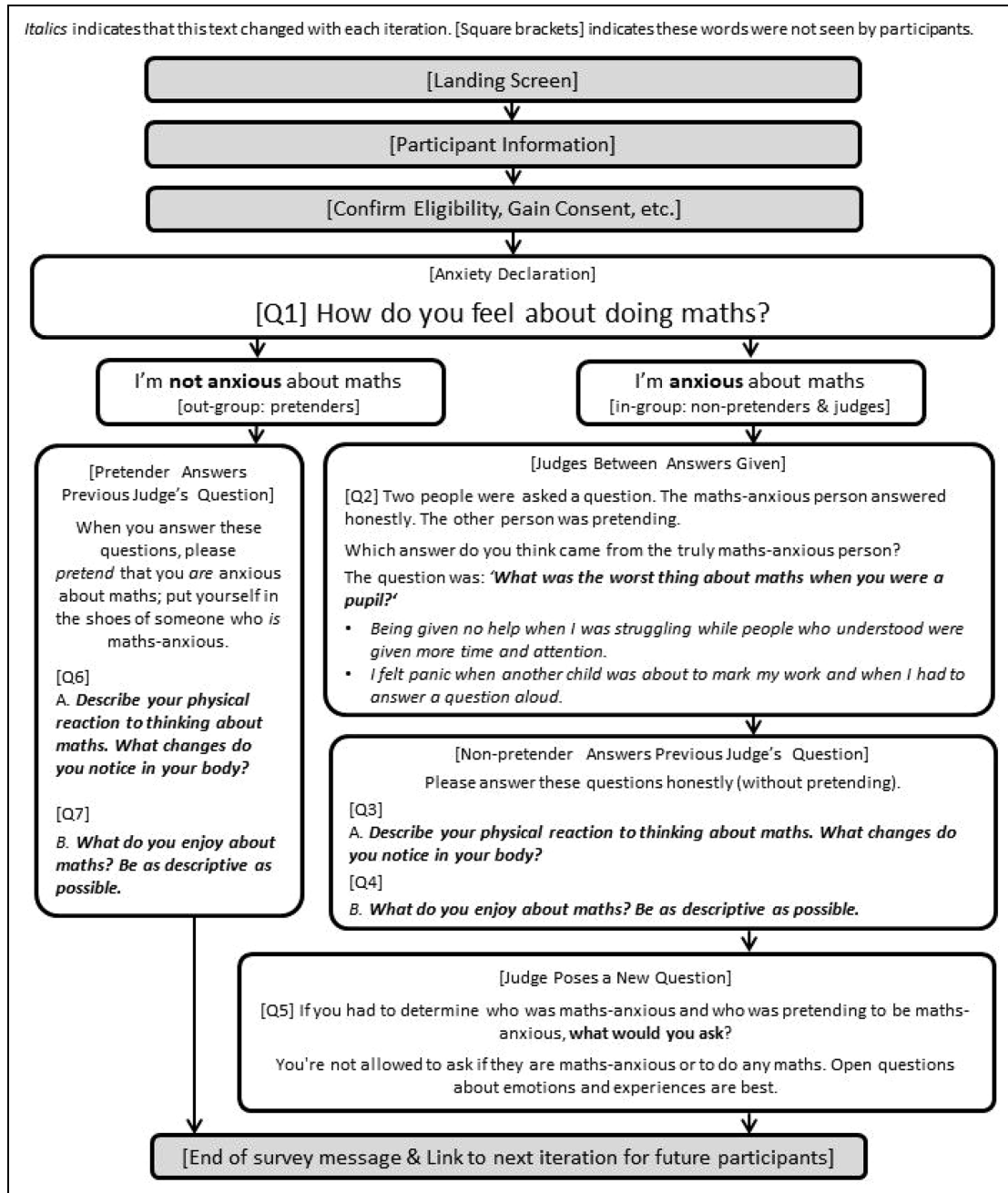


Figure 2. Item types and survey flow.

judge can see a pair of responses (one from a pretender and one from a non-pretender) to the same question.

Moving between iterations

Figure 2 shows which elements of the survey change from iteration to iteration (in bold) and also indicates that multiple versions of some items can be included. While it is not recommended that Q5 be duplicated due to the required cognitive demand, multiple versions can be used

for other item types. As with any survey, care should be taken not to over-burden participants, particularly given the demands placed on the in-group. How many questions should be included will vary from population to population depending on their anticipated engagement and therefore tolerance of the survey.

In order to give time for each new iteration to be established, recruitment was done via email in batches of 100 schools with follow-up phone calls made soon after the initial recruitment email. Most schools had not yet

forwarded the appeal to their teachers when the call was made and speaking with the researcher by phone seemed to help greatly in drawing attention to the email and encouraging recipients to share it with teaching staff. As with web surveys in general (Callegaro et al., 2015), responses were seen within a day or two of the appeal reaching teachers (once it was forwarded to them). This made it somewhat possible to control when participants completed the survey.

Each iteration was left open while its data was analysed in order to avoid frustrating any teacher attempting to complete a closed survey. How long it took to establish each new iteration depended on how many substantively new contributions had been made, but 2–3 hours was generally enough, given that the researcher was familiar with the survey platform in use and had become well-versed in moving between iterations.

In this project, an answer from each player type was selected pseudo-randomly using a random number generator, moderating in case the pretender and non-pretender answers matched and in case of one word answers (it was felt the judge would be reduced to guessing if such limited information was presented). In this project, an answer was re-selected only once. The judge-posed question asked about confidence while talking through a maths topic and the answers initially selected were, ‘not very’ and ‘not confident’. These were thought to be sufficiently similar to warrant replacing one of them. A new pseudo-random number was generated which selected, ‘Not very unless I had prepared the questions fully first’. While this was still similar it was felt that the judges now had sufficient input on which to base their judgements.

Closing the survey

Each iteration should be closed when a quota has been met. The quota could relate to the number of participants (of either type) or the number of unique answers given. In the current research, it was decided that an iteration would close when at least 10 in-group participants had answered the anxiety declaration and at least one other item. This was a pragmatic decision arising from the anticipated scarcity of maths-anxious teachers, the time it took to recruit them and the timescale for completing the project. Using a hyperlink in the closed survey message of each iteration, a link was provided so that iterations were connected in series and any participant responding with an old link would find the current iteration. A threshold for ending the survey (creating no further iterations) must also be set. In this case, the point at which judge-posed questions cease to contain new ideas (substantive saturation) was used. The current research reached substantive saturation after five iterations.

Maximising the data

The data was maximised by adopting another idea from Evans et al. (2019); transcripts. A transcript is a single judge-posed question with a corresponding pair of answers. After the close of the last survey iteration, transcripts were collated (by gathering all versions of Q2 from the various iterations) and presented to new in-group judges who were then asked to distinguish between pretender and non-pretender. These data were combined with the data captured in each iteration giving larger frequencies and a greater chance of statistical significance for each tested, judge-posed question.

Concession

In order to adapt the Game into a survey, a concession was necessary. In the original Imitation Game, judges receive a pair of responses that answer *their own* judge-posed question. In the survey version, judges are asked to pose a new question to aid discernment between players but never see its answers. Instead, they see answers to a question posed by a previous judge. This concession may have increased the difficulty of judges’ determinations since they first had to infer the authoring judge’s intentions in posing that particular question.

Ethics

As with the original game, researchers using a method such as the one discussed here, do not need to mislead participants. They can fully explain the method and therefore adult participants can give informed consent. Everyone knows that someone is pretending and there is no reason for participant-researcher trust to be harmed. Even in the case of a study topic which might provoke negative stereotypes, the role of the judge in distinguishing between pretenders and non-pretenders will eliminate the out-group pretender’s stereotype, unless the stereotype is also subscribed to by the in-group judge (in which case, one might argue that the type is valid). Participants may also use their right to withdraw, a right which should always be robustly established. Moreover, researchers may consider that for some sensitive and serious research topics (mental health, domestic abuse, and so on), a research method referred to as a game may seem insufficiently serious for some participants. Finally, it may be unwise to use either the game or the survey adaptation with minors; the iterative survey uses an unusual scenario, which, if not fully understood, could result in the validity of child/parent consent becoming questionable.

Sample

The sample of 1108 schools was drawn randomly from a sampling frame of all schools in Wales teaching children

Table 1. Judge-posed questions.

Iteration	#JPQs generated in iteration	#JPQs considered substantively new	#JPQs included in later iterations
A	9	5	3 used in B, 2 used in E
B	8	3	3 used in C
C	7	3	3 used in D
D	5	0	0
E	5	0	N/A
Total	34 ^a	11	11 (the 12th question came from the researcher)

^a14 judges who took part in iterations declined to pose a new question. Ten judges took part in transcripts and were not asked to pose new questions.

aged 3–11 years old. It included both maintained and independent schools as well as Pupil Referral Units. Schools were asked to forward the recruitment email to all relevant qualified teachers for voluntary participation making the final sample a convenience sample. Most schools (83% of those sampled) were also spoken to by phone. 182 schools participated and, over the five iterations of the survey, a total of 278 primary teachers from Wales answered at least one question. This is roughly 13.5% of all relevant schools in Wales and 2.3% of the total teacher population. 58 teachers declared MA.

Representativeness

Participants were offered a high degree of anonymity. After confirming their eligibility as a qualified teacher working in Wales in the last 2 years, the only other biographical data collected was the postcode of their school. Consequently, it is possible to assess the representativeness of the sample of schools but not of the sample of teachers. Responding schools were found to be representative of the population of schools in terms of geographical spread; the Chi Square test for goodness of fit found no statistically significant differences between sampled schools and the population of schools in each local authority. Likewise Chi Square tests found no statistically significant differences between the sample and the population in terms of language of instruction (Welsh medium, English medium, etc.), school type (primary, special, all through, etc.), funding type, or school size.

Substantive findings

The focus of this article is methodological innovation but a brief overview of the findings is necessary to demonstrate the potential of the Iterative Survey.

Judge-posed questions

A total of 34 questions were posed by maths-anxious teachers while playing in the role of the judge. Table 1 shows in which iteration questions were posed, whether they were considered substantively new and how those

questions were used. Up to three substantively new judge-posed questions were included in the following iteration or later iterations until all substantively new judge-posed questions had been presented. Iteration A included questions from the pilot study. The final iteration also included a researcher³ posed question (so that there were the same number of new questions as in other iterations).

It is thought that the number of questions posed by judges (substantively different or not, see Table 1, column 2) reduced over time, because participants were trying to pose an as yet un-posed question, correctly inferring what the researcher was looking for. However, since each iteration presented participants with more and more versions of the multiple choice item type (Q2 and Q3), the challenge of posing a novel question became harder. As the number of judge-posed questions per iteration decreased, so did the number of substantively new questions, possibly for the same reason.

It may also be that MA in teachers is a kind of ‘latent groupishness’ if maths anxious participants were drawing from a relatively limited bank of conceptions of MA. In contrast to a more actively maintained social group, such as Alcoholics Anonymous who meet regularly to discuss the experience of group membership, maths anxious teachers may not know anyone else in their group. They may not have talked to anyone about their maths-related feelings since childhood, perhaps meaning that they did not have a ready bank of vocabulary or pre-formed conscious ideas about being a maths anxious teacher from which to form questions.

Of the 12 substantively different questions posed by the judges (and in one case, the researcher), answered by both types of player, and presented to new judges, 10 lead to statistically significant differences according to a Chi Square test (see Supplemental Material). In two of those cases the majority judge vote was for the pretender answer, leaving 8 questions that resulted in the majority judge vote correctly identifying the in-group, non-pretender. The two occasions where the majority vote was for the pretender answer were as shown in Table 2.

Accepting pretender answers that win a majority of the judge vote should be considered with caution since there may be some contexts in which it is unwise. In the case of

Table 2. Judge-posed questions where the majority voted for the pretender answer.

Judge-posed question	Answers		Votes	%
Describe your physical reaction to thinking about maths. What changes do you notice in your body?	Correct (anxious non-pretender)	I get a sore and upset tummy.	5	24
	Incorrect (non-anxious pretender)	Twitchy, flushed and restless.	16	76
What happens to your emotions when you are doing maths?	Correct (anxious non-pretender)	I get frustrated when I'm confused. I feel like I should know exactly what to do and how to it.	11	29
	Incorrect (non-anxious pretender)	I feel anxious and lack confidence.	27	71

the current research, there is considerable existing MA theory against which such answers can be compared (Dowker et al., 2016). In the case of research topics where there is little existing theory or theory is not well established, it would be wise to proceed with caution rather than risk adding extraneous concepts to emerging definitions. Likewise, with highly contentious research topics, the acceptance of pretender answers may be so controversial as to damage the reputation of the research and should therefore be avoided. In contexts where individuals might be motivated and able to intentionally sabotage the research, majority voted pretender answers should not be accepted (other measures would also be needed to protect the veracity of the research).

Researcher judgement

As with any research project, researcher decision making is an indispensable part of the research process (Saunders et al., 2009). A multitude of judgements must be made about the selection of research focus, the population, the mode of recruitment and the selection of method. When using the Iterative Survey, researcher judgement is needed in determining the quota type and size that would trigger a new iteration. It is also needed in determining which judge-posed questions are considered sufficiently substantively similar that some are omitted from subsequent iterations. Judgements must be made about whether to replace randomly drawn pairs of answers in subsequent iterations on the ground of similarity. Decisions must be made about the 'seed' questions needed to populate the first iteration of a project. A number of factors including participant burden, expected sample size, available budget and timescale must be balanced in deciding how granular these judgements should be.

An example of researcher judgement used here was in deciding whether, 'What do you enjoy about maths?' overlapped sufficiently with, 'Why do you enjoy maths?' to test only the former (while linguistically different, both questions would evoke similar answer content). Conversely, 'How do you feel when asked to do mental

maths on the spot?' was judged to be sufficiently different to, 'How do you feel when you're faced with word problems in maths?' that the latter was tested as a separate judge-posed question.

Discussion

Does the Iterative Survey retain the Game's ability to capture 'groupishness'?

The Iterative Survey appears to be able to capture groupishness. A simpler enquiry, such as, 'What is it like to feel anxious about maths?' might adequately capture elements of MA groupishness but it would not be evident which are idiosyncratic and which are shared yet distinctive. By framing the enquiry as in Q5, an extra level of information is captured, as in the Liar Paradox (see above). Not only are elements of MA captured, their degree of groupishness is quantified. In this case at least, the Iterative Survey has captured the groupishness of primary school teachers' experience of MA.

Generalisation

As above, the current sample is representative of the population at the school level on the five factors analysed. Given the fact that maths anxious teachers are hidden within the wider teacher population, it is thought that the sample size is respectable. These combine to indicate that conclusions can be cautiously generalised.

Williams (2000) contends that it is inevitable and appropriate that 'moderatum' generalisations be drawn from interpretive research in order that research might 'say something of something' (Geertz, 2000). Such (speculative) generalisations, made on the basis of inferences about regularity in each setting, state that aspects of a specific instance could be considered to represent a broader set and are the type of generalisations made in everyday life (Williams, 2000). Williams also states that moderatum generalisations made to closely aligned instances have greater validity than those made to less similar instances. Therefore, speculative moderatum generalisations were

made and could be made following future uses of the Iterative Survey.

Further, a strength of the Iterative Survey is that it is more easily scalable than the original Imitation Game, making the use of large sample sizes, with their inherent advantages for population representativeness, more easily attainable. While the sample used in this research on a hidden population was modest, any future use of the Iterative Survey that reaches substantive saturation and uses a representative sample will support confident generalisations being made from the data. Scaling the sample size up will also produce a greater degree of confidence and sharpness in the generalisations possible.

The project reported in this paper was limited by budget and timescale, leading to the inception of the Iterative Survey, but future projects with fewer restrictions could make a direct evaluation of the robustness of the Iterative Survey by close comparison between Imitation Games and Iterative Surveys conducted on the same topic. If a set of Imitation Games were played with primary teachers in Wales, exploring MA, then the findings could be compared with the findings reported here. Or Iterative Surveys could be conducted on topics previously explored by full Imitation Games and their conclusions compared. This thought is discussed further in Future Development, below.

Limitations

Despite the survey's advantages, limitations remain. As with the Game, the cognitive challenges involved in the judge and pretender roles are considerable. The pretender must draw on their ability to empathise, and imagine their answer to the judge-posed question as if they were maths-anxious. This demands both self- and other-awareness (Gallup, 1996). The judge also has considerable cognitive challenges, both in posing a new question and in distinguishing pretender from non-pretender. They must consider both the way that a non-pretender would answer honestly, and the way a pretender might attempt to mimic the non-pretender in order to create a question that might result in identifying the pretender. While when acting as the non-pretender, participants have the advantage of answering as themselves, they must still demonstrate a high degree of self-awareness by sorting through their topic-related experiences, thoughts and feelings to select those that they feel are most shared and distinctive.

As with all forms of research, participants are balancing their motivation to participate with the time-cost of participation (Callegaro et al., 2015), alongside considerations such as self-monitoring to stay within the instructions of the method and give what they perceive to be socially acceptable answers (Dodou and de Winter, 2014). In the original Game, the presence of other players in the same room at the same time may provide more

motivation through enjoyment of game play. Since participants do not know the identities of the players they are interacting with, there may be an air of mystery that may engage participants and increase their motivation in comparison with the survey version where the other players are further removed.

Both the Game and the survey demand considerable literacy skills. While much of the instruction-giving for the Game can be done through verbal instructions, playing the Game involves reading contributions from the other two players (in the role of the judge) and writing one's own contributions clearly and succinctly. The Iterative Survey involves less time-pressure since there are no other players awaiting your contributions. However the relatively lengthy instructions (compared with most surveys) demand good readings skills which are essential in order to understand the novel scenario.

Consequently, the cognitive, digital and literacy abilities of the study population must be carefully considered before use of either the Game or the Iterative Survey. Additionally, as with any web survey, the population must also have excellent internet coverage to enable the capture of good quality data.

Future development

There is considerable scope and need for future development and testing of the method.

Evans et al. (2023) have recently published the results of a replication study comparing the outcomes of Imitation Games about religion with the pre-existing findings of large-scale panel survey data. There is also scope to test the robustness of the Iterative Survey method by application to a wider range of topics, particularly those with which the original Game has already been played (see Origins of the Imitation Game, above). Using the survey with, for example, Coeliac patients and their dieticians, would allow close comparison of the replicability of findings and the methods' robustness.

The Iterative Survey could also be used with groups who might be difficult to recruit for synchronous Games, for example patients who are housebound and their General Practitioners, incarcerated youth offenders and those involved in their rehabilitation, missionaries working abroad and the charities who fund them (provided that such populations meet the cognitive, literacy and computer requirements mentioned above).

The survey could also facilitate explorations with groups for whom Games would be impossible for other reasons. Returning to MA, let us imagine we want to know how well teachers understand the MA experience of their young pupils. Playing synchronous Games would be problematic since pupil answers may be easily distinguished from teacher answers if the tone of voice and quality of spelling and grammar gave them away. The

survey version provides an opportunity for pupil contributions to be edited to sound more adult-like or vice versa.

In future versions of the Iterative Survey, it may also be helpful to adopt another element of the original Game; recording the judge's level of confidence that they have correctly identified the non-pretender (as happens in Games). Greater nuance could be brought to calculations regarding the ability of judge-posed questions to identify capture groupishness by offering, say a four-point scale on which judges can rate their confidence. For example, when evaluating the questions, determinations made with low confidence could be disregarded, or the proportion of unsure judges could be quantified in order to provide a separate metric of the question's utility.

While the Imitation Game itself remains a novel social science research method, the Iterative Survey version is clearly in its infancy and presents many opportunities for development.

Conclusion

The current research reports on the adaptation of the Imitation Game into Iterative Survey form, moving from a live synchronous Game to an asynchronous survey, while maintaining the method's ability to identify and assess shared and distinctive elements of group membership. Reported here, is its first use, exploring the maths anxiety of primary teachers in Wales.

The Iterative Survey is potentially an important method for the exploration of group identity. Through its ability to analyse distinctive elements of group membership, the Iterative Survey offers researchers a scalable way to capture group members' own conceptions of group identity, accessing potentially rich data about groups, previously only obtainable through qualitative methods. The adaptation to Iterative Survey not only reduces participant burden but also offers researchers on tight budgets and timescales the opportunity to give voice and audience to a variety of group members, including those who might otherwise not be heard. It is robust, even when faced by someone falsely claiming group membership; they will be found out and their contributions disregarded, unless they demonstrate such effective understanding of group membership that they are able to fool genuine members, in which case, one might argue, their contributions are as useful as true group members. The Iterative Survey offers a unique contribution to the social science of group exploration.

The potential applications of the Iterative Survey are manifold, both in education research and beyond. The Iterative Survey could shed light on how well non-dyslexic teachers understand their dyslexic pupils, how well school inspectors understand the experience of being a classroom

practitioner, or how effectively head teachers can put themselves in the shoes of, say, newly qualified teachers, all of relevance for the professional development of teachers and the improvement of accountability and assessment systems. While there might be practical recruitment problems with playing a live Game between fraudsters and police officers, the Iterative Survey might allow insight into the effectiveness of police understandings of fraud, perhaps of benefit in the training of officers. Similarly, playing a live Game between foster parents and children in foster care might be problematic due to the differing use of grammar and tone, the Iterative Survey could identify distinctive and shared elements of the foster care experience which might be of use to the training of new foster parents. Returning to the Game's first use in the medical world, the Iterative Survey opens a wider range of options; it could explore how well do paramedics (who might be difficult to recruit for a synchronous Game) understand the emergency-scenario medical needs of people with, say, autoimmune conditions.

While the Iterative Survey is in its infancy, its further development and wider use promises access to valuable insights in any context where people form groups.

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author received financial support for the research from UKRI. Information on the data underpinning this publication, including access details, can be found in the Cardiff University Research Data Repository.

Ethical approval

This research has been given a favourable decision by the Cardiff School of Social Sciences Ethics committee (SREC 327).

ORCID iD

Fay Cosgrove  <https://orcid.org/0009-0000-2713-3442>

Supplemental material

Supplemental material for this article is available online.

Notes

1. In Wales, a primary teacher typically teaches a class of around 30 pupils for the whole academic year. Pupils begin statutory education at age 5 and leave primary school at age 11. Teachers usually teach all subjects of the curriculum (Nicholls, 2004).

2. In the Turing Test, a human judge asks a question of a human and a computer, receiving the answers anonymously. If the computer can fool the human judge into selecting the computer as the true human, it is considered to have artificial intelligence.
3. The researcher was, at the time of Iteration E, eligible to be a participant having been employed to teach primary maths in Wales in the last 2 years.

References

- Arminen I, Segersven O and Simonen M (2019) Active and latent social groups and their interactional expertise. *Acta Sociologica* 62(4): 391–405.
- Artemenko C, Masson N, Georges C, et al. (2021) Not all elementary school teachers are scared of math. *Journal of Numerical Cognition* 7(3): 275–294.
- Ashcraft M, Krause J and Hopko D (2007) Is math anxiety a mathematical learning disability? In: Berch D and Mazzocco M (eds) *Why Is Math So Hard for Some Children? The Nature and Origins of Mathematical Learning Difficulties and Disabilities*. (pp. 329–348) Baltimore: Paul Brookes Publishing.
- Beilock SL, Gunderson EA, Ramirez G, et al. (2010) Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences USA* 107(5): 1860–1863.
- Boyce H (2009) *Interactional expertise in practice: Researching dietitians' knowledge of living with coeliac disease*. Doctoral dissertation, Cardiff University.
- Buckley S and Sullivan P (2023) Reframing anxiety and uncertainty in the mathematics classroom. *Mathematics Education Research Journal* 35: 157–170.
- Callegaro M, Manfreda K and Vehovar V (2015) *Web Survey Methodology*. London: Sage.
- Chinn S (2020) *More Trouble With Maths: A Complete Manual to Identifying and Diagnosing Mathematical Difficulties*, 3rd edn. London: Routledge.
- Collins H (2016) *Gravity's Kiss: The Detection of Gravitational Waves*. Cambridge, MA: MIT Press.
- Collins H and Evans R (2017) Probes, surveys, and the ontology of the social. *Journal of Mixed Methods Research* 11(3): 328–341.
- Collins H, Evans R, Hall M, et al. (2019) Bonfire night and burns night: Using the imitation game to research English and Scottish identities. In: Caudill DS., Conley SN, Gorman ME, et al. (eds) *The Third Wave in Science and Technology Studies*. Cham: Palgrave Macmillan, pp.109–131.
- Collins H, Evans R, Ribeiro R, et al. (2006) Experiments with interactional expertise. *Studies in History and Philosophy of Science Part A* 37(4): 656–674.
- Dodou D and de Winter J (2014) Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior* 36: 487–495.
- Dowker A, Sarkar A and Looi C (2016) Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology* 7: 508.
- Evans R (2023) *Personal Interview*, 3 November.
- Evans R, Collins H, Weinel M, et al. (2019) Groups and individuals: Conformity and diversity in the performance of gendered identities. *The British Journal of Sociology* 70(4): 1561–1581.
- Evans R, Collins H, Weinel M, et al. (2023) Evaluating the imitation game as a method for comparative research: A replication study using imitation games about religion. *International Journal of Social Research Methodology* 26(3): 249–262.
- Evans R and Crocker H (2013) The Imitation Game as a method for exploring knowledge (s) of chronic illness. *Methodological Innovations Online* 8(1): 34–52.
- Fernández-Sogorb A, Vicent M, González C, et al. (2020) Attributional style in mathematics across anxiety profiles in Spanish children. *Sustainability* 12(3): 1173.
- Foley A, Herts J, Borgonovi F, et al. (2017) The math anxiety-performance link: A global phenomenon. *Current Directions in Psychological Science* 26(1): 52–58.
- Gallup (1996) Self-awareness and the evolution of social intelligence. *Behavioural Processes* 42(1998): 239–247.
- Geertz, C. (2000). Deep Play: Notes on the Balinese Cockfight. In: Crothers, L., Lockhart, C. (eds) *Culture and Politics*. Palgrave Macmillan, New York.
- Hembree R (1990) The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education* 21(1): 33–46.
- Hill F, Mammarella I, Devine A, et al. (2016) Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences* 48: 45–53.
- Maloney E, Ramirez G, Gunderson E, et al. (2015) Intergenerational effects of parents' math anxiety on children's math achievement and anxiety. *Psychological Science* 26(9): 1480–1488.
- Miniukovich A, De Angeli A, Sulpizio S, et al. (2017) Design guidelines for web readability. In: *Proceedings of the 2017 Conference on Designing Interactive Systems*, Edinburgh, pp.285–296. New York, NY: Association for Computing Machinery.
- Nicholls G (2004) *An Introduction to Teaching: A Handbook for Primary and Secondary School Teachers*. London: Routledge.
- Organisation for Economic Co-operation and Development (2015) Pisa 2015 Database. Available at <https://www.oecd.org/pisa/data/> (accessed 12 August 2022).
- Parsons C (1974) The liar paradox. *Journal of Philosophical Logic* 3(4): 381–412.
- Richardson F and Suinn R (1972) Mathematics anxiety rating scale: Psychometric data. *Journal of Counselling Psychology* 19: 551–554.
- Saunders M, Lewis P and Thornhill A (2009) *Research Methods for Business Students*. New York, NY: Pearson Education.
- Schaeffer M, Rozek C, Maloney E, et al. (2020) Elementary school teachers' math anxiety and students' math learning: A large-scale replication. *Developmental Science* 24(4): e13080.
- Segersven O, Arminen I and Simonen M (2024) Acculturation among Finnish Somalis: An imitation game inquiry into bicultural fluency. *International Migration Review* 58(2): 680–705.
- Segersven O and Heino A (2019) *Email to Anonymous*. 19 March.
- Turing AM (1950) Computing machinery and intelligence. *Mind* 59: 433–460.
- Wang Z, Lukowski SL, Hart SA, et al. (2015) Is math anxiety always bad for math learning? The role of math motivation. *Psychological Science* 26(12): 1863–1876.

Wehrens R and Walters B (2018) Understanding each other in the medical encounter: Exploring therapists' and patients' understanding of each other's experiential knowledge through the Imitation Game. *Health* 22(6): 558–579.

Welsh Government (2022). National survey for Wales. Available at: <https://www.gov.wales/national-survey-wales> (accessed 9 April 2024).

Williams M (2000) Interpretivism and generalisation. *Sociology* 34(2): 209–224.

Yucedag-Ozcan A and Brewer S (2011) Adaptation of Mathematics Anxiety Rating Scale-revised (MARS-R) for adult online students. Available at: https://www.researchgate.net/publication/266180761_Adaptation_of_Mathematics_Anxiety_Rating_Scale-Revised_MARS-R_for_Adult_Online_Students (accessed 20 September 2024).

Author biographies

Fay Cosgrove is an accomplished primary school teacher with 18 years of experience, working regionally and nationally to improve maths outcomes through continuing professional development (CPD) and assessment. She is now a PhD researcher developing a bilingual psychometric scale to measure maths anxiety in primary teachers in Wales. She also works with an educational think tank, Learnus, conducting and co-ordinating research into teacher attitudes to maths. Her research interests include teacher and pupil attitudes to maths, maths CPD, and teacher self-reflection.

Appendix A

Search terms used to seek examples of a social science game being converted into a survey;

Online university library search

'game' AND 'survey' OR 'questionnaire' AND 'convert' OR 'adapt' OR 'transform' OR 'modify' in any field

In subject areas: Social Sciences, Education & Educational Research, Psychology, Polls & Surveys, Game Theory, Surveys & Questionnaires, Gamification

Language: English

All item types

105 results

Google scholar search

All of the words: game survey

With at least one of the words: convert adapt, transform, modify

Anywhere in the article

Read titles (and some abstracts) of first 100 results

Google scholar search

All of the words: game questionnaire

With at least one of the words: convert adapt, transform, modify

Anywhere in the article

Read titles (and some abstracts) of first 100 results