A ten-minute presentation of my research interests

Matthew Inglis

Loughborough University

Centre for Mathematical Cognition

My Interests

- Who am I?

My Interests

- Who am I?
- Interested in **mathematical cognition**

My Interests

- Who am I?
- Interested in **mathematical cognition**
  - I conceive this to mean the ways in which mathematical information is processed

My Interests

- Who am I?
- Interested in **mathematical cognition**
    - I conceive this to mean the ways in which mathematical information is processed
    - Often, but not always, this is in an educational context

My Interests

- Who am I?
- Interested in **mathematical cognition**
  - I conceive this to mean the ways in which mathematical information is processed
  - Often, but not always, this is in an educational context
  - An interdisciplinary endeavour involving psychologists, educators, neuroscientists (and perhaps some philosophers and anthropologists).

Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

"Distinctively, we aim to bridge the gap between basic research on mathematical cognition and more applied work that involves designing and evaluating research-informed pedagogical materials. Some examples of our recent work:

Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

"Distinctively, we aim to bridge the gap between basic research on mathematical cognition and more applied work that involves designing and evaluating research-informed pedagogical materials. Some examples of our recent work:

- Theresa Wege and colleagues investigated how young children develop the ability to count abstract units;

Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

"Distinctively, we aim to bridge the gap between basic research on mathematical cognition and more applied work that involves designing and evaluating research-informed pedagogical materials. Some examples of our recent work:

- Theresa Wege and colleagues investigated how young children develop the ability to count abstract units;
- Francesco Sella and colleagues studied the influence that choosing to study mathematics at post-compulsory levels has on brain development;

Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

"Distinctively, we aim to bridge the gap between basic research on mathematical cognition and more applied work that involves designing and evaluating research-informed pedagogical materials. Some examples of our recent work:

- Theresa Wege and colleagues investigated how young children develop the ability to count abstract units;
- Francesco Sella and colleagues studied the influence that choosing to study mathematics at post-compulsory levels has on brain development;
- Colin Foster and colleagues are harnessing basic research insights to develop a complete, fully resourced, and free-to-access mathematics curriculum;

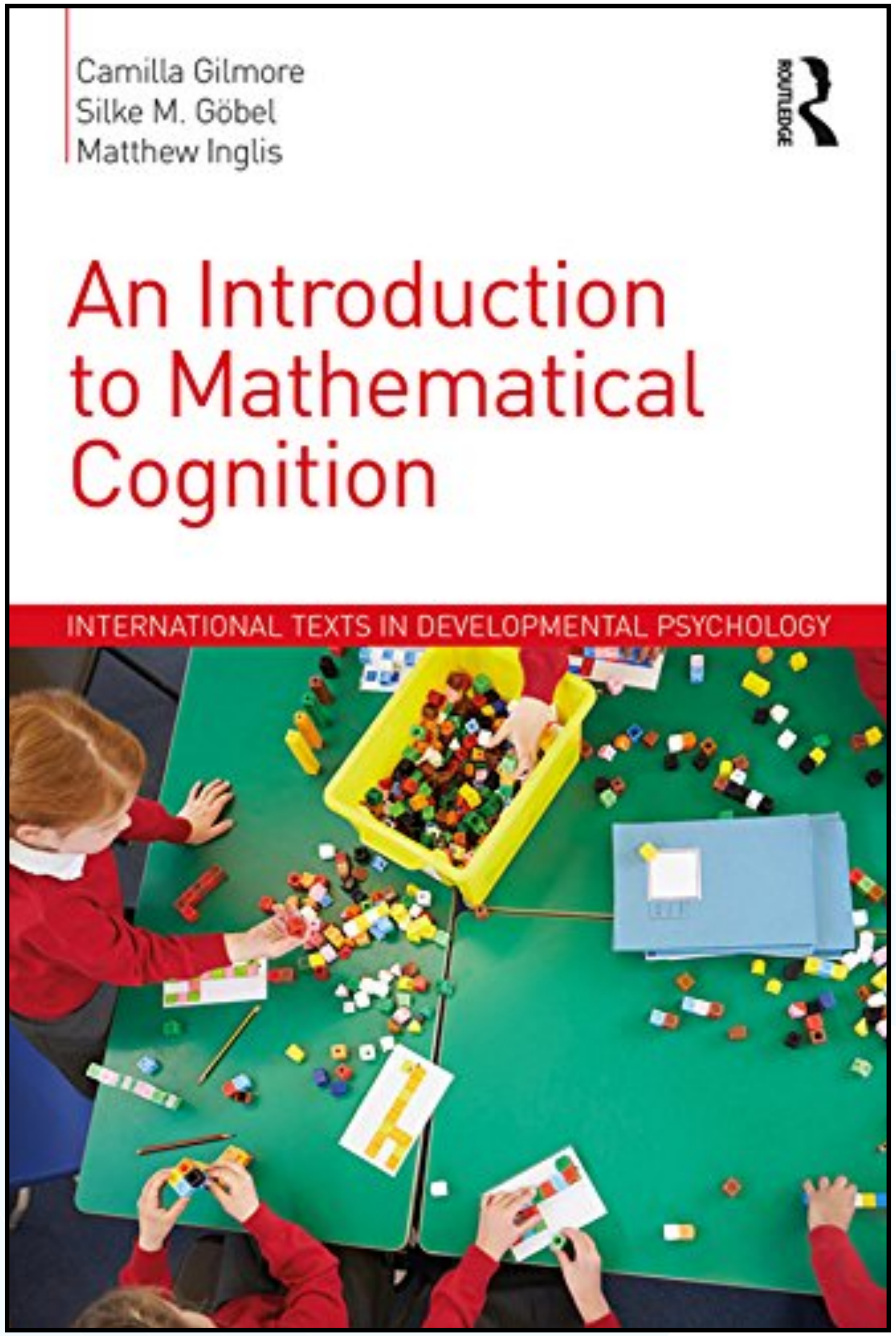
Mathematical Cognition

I direct Loughborough's Centre for Mathematical Cognition: www.cmc.ac.uk

"Distinctively, we aim to bridge the gap between basic research on mathematical cognition and more applied work that involves designing and evaluating research-informed pedagogical materials. Some examples of our recent work:

- Theresa Wege and colleagues investigated how young children develop the ability to count abstract units;
- Francesco Sella and colleagues studied the influence that choosing to study mathematics at post-compulsory levels has on brain development;
- Colin Foster and colleagues are harnessing basic research insights to develop a complete, fully resourced, and free-to-access mathematics curriculum;
- Hugues Lortie-Forgues and Matthew Inglis studied how educational interventions are currently evaluated, arguing that existing methods typically provide uninformative results and suggesting how the situation could be improved."

Mathematical Cognition



- Nonsymbolic number
- Symbolic number
- Development of arithmetic skills
- Understanding of arithmetic concepts (e.g. commutativity, inversion, multiplicative reasoning), conceptual and procedural knowledge
- Individual differences (e.g., dyscalculia, mathematics anxiety)
- Number systems
- Algebra and equivalence
- Mathematical argumentation and proof
- Logic, conditional reasoning and mathematics

Theresa's Work

Theresa Wege's PhD:
How we think about numbers: Early counting and mathematical abstraction

Worked with typically developing four and five year old children.

Theresa’s Work

Journal of Experimental Child Psychology 225 (2023) 105533

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Counting many as one: Young children can understand sets as units except when counting

Theresa Elise Wege [a,*], Bert De Smedt [b], Camilla Gilmore [a], Matthew Inglis [a]

[a] Centre for Mathematical Cognition, Loughborough University, Loughborough LE11 3TU, UK
[b] Parenting and Special Education Unit, Katholieke Universiteit (KU) Leuven, B-3000 Leuven, Belgium

A R T I C L E I N F O

Article history:
Received 15 November 2021

A B S T R A C T

Young children frequently make a peculiar counting mistake.

E: What kinds of animals are there?

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: How many blocks are there?

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: How many blocks are there?
C: Four

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

E: Please sort the animals so that the kinds are together
C: [Sorts]

E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: How many blocks are there?
C: Four

E: Remember, each kind of animal has one block, how many kinds of animals are there?

E: What kinds of animals are there?
C: Sheep, pig, cows, horses

E: How many kinds of animals are there?
C: Nine

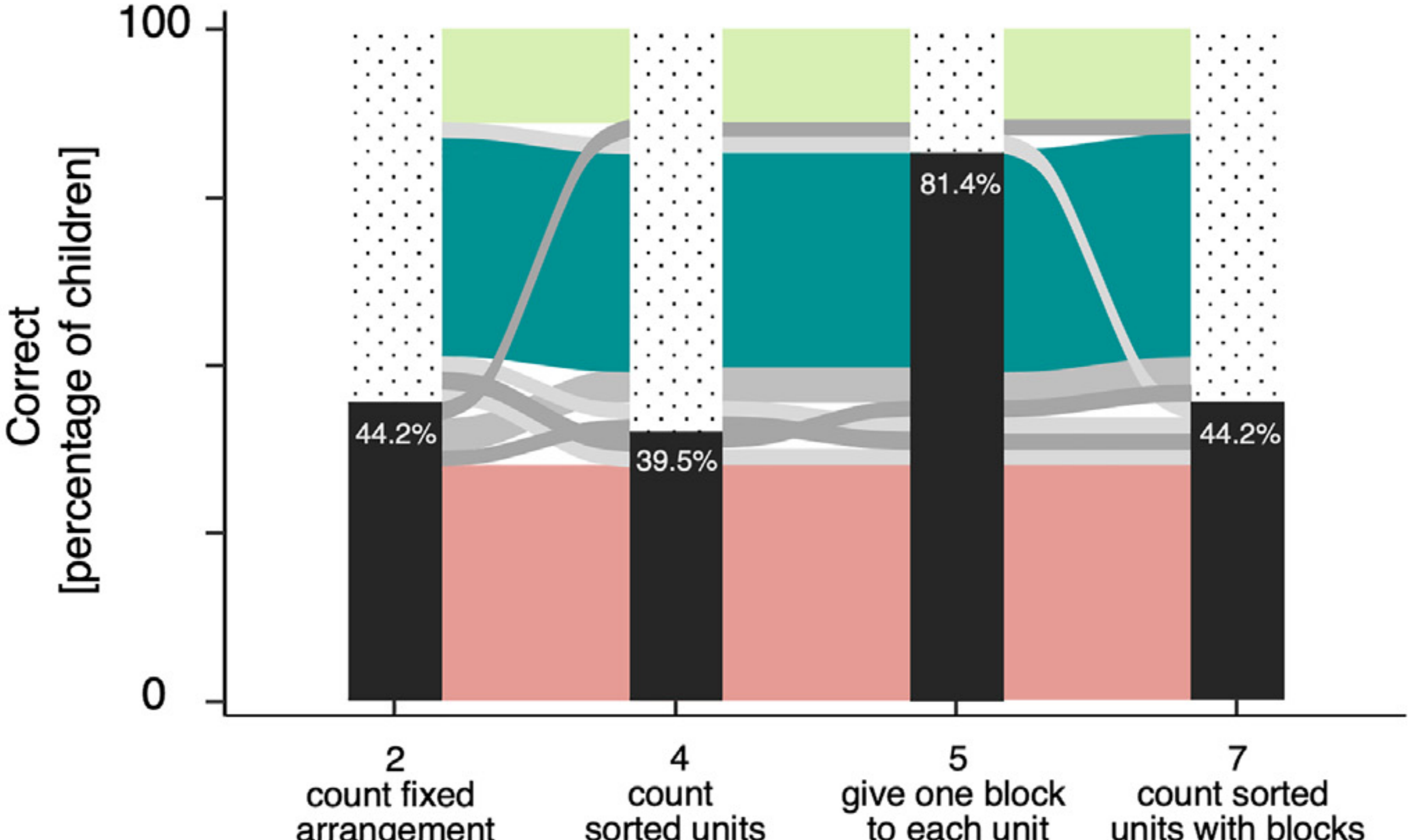
E: Please sort the animals so that the kinds are together
C: [Sorts]

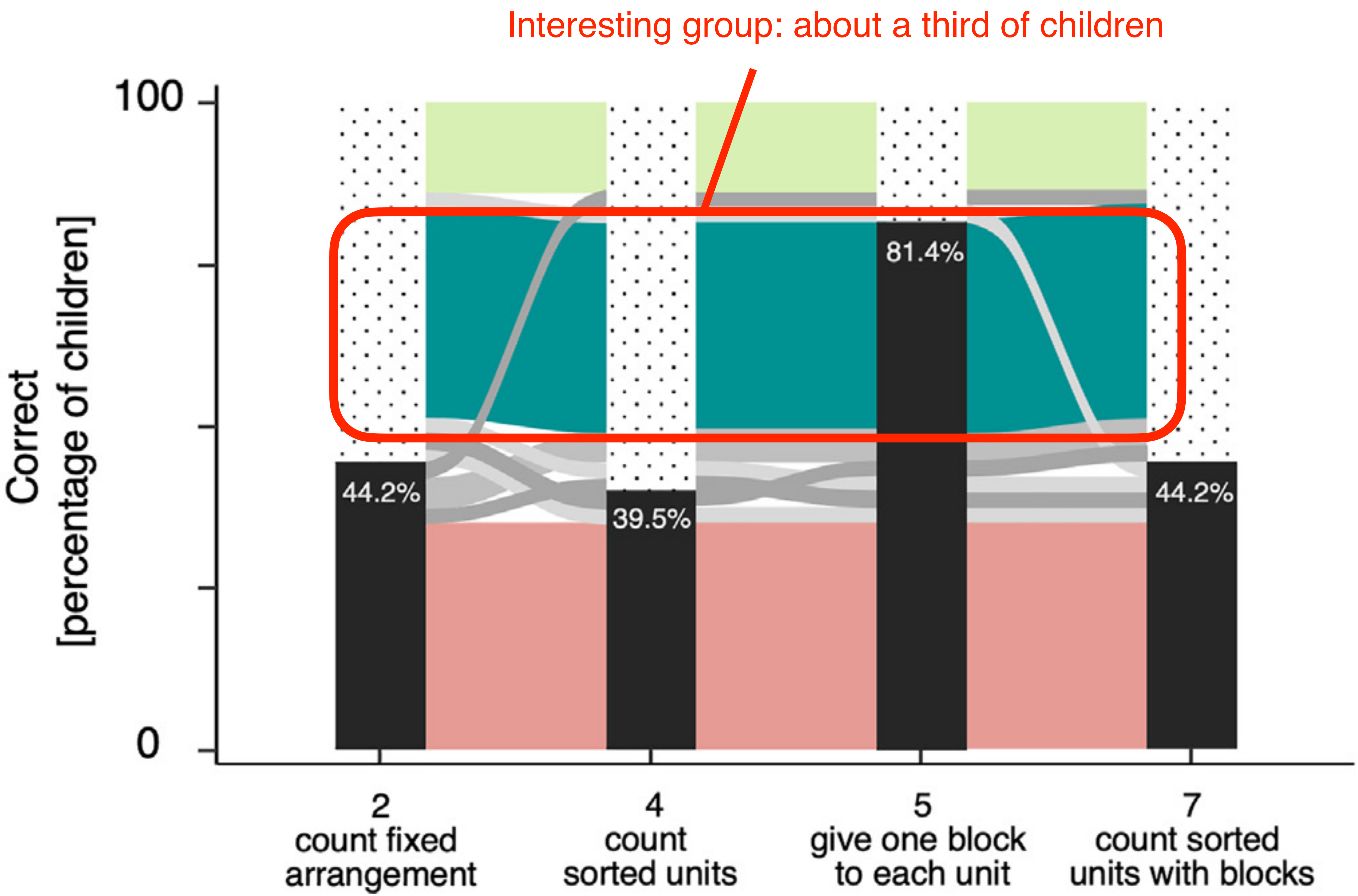
E: We now have groups of the kinds of animals. How many kinds of animals are there?
C: Nine

E: Please give a block to each kind of animal
C: Gives blocks

E: How many blocks are there?
C: Four

E: Remember, each kind of animal has one block, how many kinds of animals are there?
C: Nine


100
0
Correct
[percentage of children]
44.2%
39.5%
81.4%
44.2%
2
count fixed
arrangement
4
count
sorted units
5
give one block
to each unit
7
count sorted
units with blocks


Interesting group: about a third of children
100
0
Correct
[percentage of children]
44.2%
39.5%
81.4%
44.2%
2
count fixed
arrangement
4
count
sorted units
5
give one block
to each unit
7
count sorted
units with blocks

Conclusions

- In contrast to previous assumptions, unitizing is not sufficient for counting (at least in the context of abstract units like “kinds of animals” or “colour”).

Conclusions

- In contrast to previous assumptions, unitizing is not sufficient for counting (at least in the context of abstract units like “kinds of animals” or “colour”).
- Children can name and sort abstract units, and create one-to-one correspondences with them, without being able to count abstract units.

Conclusions

- In contrast to previous assumptions, unitizing is not sufficient for counting (at least in the context of abstract units like “kinds of animals” or “colour”).
- Children can name and sort abstract units, and create one-to-one correspondences with them, without being able to count abstract units.
- Main theoretical conclusion: Gelman & Gallistel’s (1978) abstraction principle (anything can be counted) is a non-trivial developmental achievement.

Conclusions

- In contrast to previous assumptions, unitizing is not sufficient for counting (at least in the context of abstract units like “kinds of animals” or “colour”).
- Children can name and sort abstract units, and create one-to-one correspondences with them, without being able to count abstract units.
- Main theoretical conclusion: Gelman & Gallistel’s (1978) abstraction principle (anything can be counted) is a non-trivial developmental achievement.
- Next question: How can we facilitate it’s development?

87–109
10.1093/philmat/nku014
Philosophia Mathematica
Advance Access publication July 25, 2014

Beauty Is Not Simplicity: An Analysis of Mathematicians' Proof Appraisals[†]

Matthew Inglis* and Andrew Aberdein**

*Mathematics Education Centre, Loughborough University, U.K.
E-mail: m.j.inglis@lboro.ac.uk
**School of Arts and Communication, Florida Institute of Technology, U.S.A.
E-mail: aberdein@fit.edu

ABSTRACT

What do mathematicians mean when they use terms such as 'deep', 'elegant', and 'beautiful'? By applying empirical methods developed by social psychologists, we demonstrate that mathematicians' appraisals of proofs vary on four dimensions: aesthetics, intricacy, utility, and precision. We pay particular attention to mathematical beauty and show that, contrary to the classical view, beauty and simplicity are almost entirely unrelated in mathematics.

1. INTRODUCTION

Mathematical conversations are full of value judgements. Mathematicians talk of 'beautiful', 'deep', 'insightful', and 'interesting' proofs, and award each other prizes on the basis of these assessments. Validity or applicability are almost never the decisive criteria for such awards. Instead the citations for mathematical prizes are full of aesthetic judgements: nine of the eleven Abel Prize citations since its foundation have characterised the prizewinner or their work as 'deep', and the work of the remaining two was lauded for its beauty and ingenuity [Holden and Piene, 2009; 2013]. Furthermore, many of the most eminent researchers have suggested that it is these value judgements which drive their research agendas. Hermann Weyl even claimed to prioritise beauty over

[†]We are extremely grateful to Lara Alcock, Donald Gillies, and Dirk Schlimm for providing insightful comments on earlier versions of this work. Early drafts of this paper were presented at the Loughborough Proof Reading Workshop (2013), the Mathematical Cultures Research Network (London, 2013), the Second International Meeting of the Association for the Philosophy of Mathematical Practice (Urbana-Champaign, 2013), and the Rutgers Proof Comprehension Workshop (2014), and we thank the audiences for their valuable remarks.
This work was supported by a Royal Society Worshipful Company of Actuaries Research Fellowship to MI. AA is grateful to Florida Institute of Technology for granting sabbatical leave.

Research Paper

PAIN®

People in pain make poorer decisions

Nina Attridge[a,*], Jayne Pickering[a], Matthew Inglis[a], Edmund Keogh[b], Christopher Eccleston[b,c]

Abstract

Chronic pain affects 1 in 5 people and has been shown to disrupt attention. Here, we investigated whether pain disrupts everyday decision making. In study 1, 1322 participants completed 2 tasks online: a shopping-decisions task and a measure of decision outcomes over the previous 10 years. Participants who were in pain during the study made more errors on the shopping task than those who were pain-free. Participants with a recurrent pain condition reported more negative outcomes from their past decisions than those without recurrent pain. In study 2, 44 healthy participants completed the shopping-decisions task with and without experimentally induced pain. Participants made more errors while in pain than while pain-free. We suggest that the disruptive effect of pain on attending translates into poorer decisions in more complex and ecologically valid contexts, that the effect is causal, and that the consequences are not only attentional but also financial.

Keywords: Pain, Cognitive disruption, Decision making, Finances, Numeracy

1. Introduction

The disruptive effect of pain on attending has been demonstrated with experimentally induced pain,[30,40] chronic pain,[9,15] and transient pain such as headache.[4,24,31] This field has predominantly focused on simple cognitive processes (although sometimes using complex tasks combining multiple executive functions[25]). Few studies have examined the effects of pain on higher-level cognition. One which did found that clinical pain was associated with less abstract thinking,[20] whereas another found no evidence that experimentally induced pain affected abstract thinking.[2] Here, we focus on the potential impact of pain on higher-level real-world cognitive tasks requiring attention, namely numerical reasoning and decision making, which have serious consequences if one gets them wrong.

Reasoning and decision making are required in many areas of life and are influenced by various cognitive and emotional factors. Here, we focus on numeracy as a domain that is important in many areas of life, including budgeting, choosing a mortgage, and choosing insurance plans. Despite its importance, numeracy in adults is poor. In the quantitative domain of the USA's 2003 National Assessment of Adult Literacy, 55% of adults performed at a basic or below basic level (at best being able to locate easily identifiable quantitative information and solve one-step arithmetic problems when the operation was specified or easily inferred).[26] These findings were echoed in a 2016 UK Money Advice Service study[29] into the public's ability to choose the best supermarket deals for 4 products. Although 74% of participants chose the best deal for at least one product, only 2% chose optimally for all 4.

Attention is important for learning and performing numerical operations in both children and adults.[13,36] Attention is also important in decision making, where we need to consider various options, estimate their likely outcomes, and then hold these in mind while choosing among them. Given that pain impairs attention, it may also influence numerical decision making. Indeed, there is some initial evidence that this is the case. Placing a hand into ice-cold water changed participants' risk-taking on a financial decision-making task.[35]

We investigated the effect of pain on everyday decision making. In study 1, a large general population sample recruited online reported whether they were currently in pain and whether they had any recurrent pain conditions. They completed 2 tasks: the shopping-decision task used by the Money Advice Service[29] and the Decision Outcomes Inventory (DOI[10]), which measures real-world outcomes of everyday decisions made over the previous 10 years. We hypothesized that participants who were in pain would find the best shopping deal on fewer items than participants who were pain-free. If the effect of pain on attention does translate into poorer decision making, the outcomes of these poor decisions may accumulate in people with chronic pain. We therefore hypothesized that participants with pain that had lasted for 3 months or longer would report more negative decision outcomes on the DOI than other participants. In study 2, we took an experimental approach to determine a causal relationship: participants completed an extended shopping-decisions task with their hand in warm or painfully cold water.

2. Study 1 method

2.1. Design and procedure

Participants (N = 1322) took part online and were recruited via Amazon's Mechanical Turk (N = 658) and Prolific.ac (N = 664). Research has shown data collected online for psychology studies are reliable[11,33] and that samples tend to be more diverse than traditional university-based samples.[27] The large

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] Mathematics Education Centre, Loughborough University, Loughborough, United Kingdom, [b] Centre for Pain Research, University of Bath, Bath, United Kingdom, [c] Department of Clinical and Health Psychology, Ghent University, Ghent, Belgium

**Corresponding author. Address: 0.28 Schofield Building, Mathematics Education Centre, Loughborough University, Loughborough, Leicestershire, LE11 3TU, United Kingdom. Tel.: +441509 223315. E-mail address: n.f.attridge@lboro.ac.uk (N. Attridge).*

PAIN 160 (2019) 1662–1669

http://dx.doi.org/10.1097/j.pain.0000000000001542

Synthese (2021) 198 (Suppl 26):S6369–S6392
https://doi.org/10.1007/s11229-019-02234-5

S.I. : ENABLING MATHEMATICAL CULTURES

Functional explanation in mathematics

Matthew Inglis[1] · **Juan Pablo Mejía-Ramos**[2]

Received: 31 May 2018 / Accepted: 25 April 2019 / Published online: 22 May 2019

Abstract
Mathematical explanations are poorly understood. Although mathematicians seem to regularly suggest that some proofs are explanatory whereas others are not, none of the philosophical accounts of what such claims mean has become widely accepted. In this paper we explore Wilkenfeld's (Synthese 191:3367–3391, 2014) suggestion that explanations are those sorts of things that (in the right circumstances, and in the right manner) generate understanding. By considering a basic model of human cognitive architecture, we suggest that existing accounts of mathematical explanation are all derivable consequences of Wilkenfeld's 'functional explanation' proposal. We therefore argue that the explanatory criteria offered by earlier accounts can all be thought of as features that make it more likely that a mathematical proof will generate understanding. On the functional account, features such as characterising properties, unification, and salience correlate with explanatoriness, but they do not define explanatoriness.

Keywords Explanation · Mathematics · Mathematical practice · Understanding

What are mathematical explanations? This question has generated substantial interest among philosophers. A number of competing accounts of mathematical explanation have been proposed (e.g., Kitcher 1981; Lange 2014; Steiner 1978), but all have well-established limitations. Our primary goal in this paper is to explore the consequences for mathematics of Wilkenfeld's (2014) notion of *functional explanation*. Roughly speaking, Wilkenfeld suggested that explanations are simply those things that, in an appropriate manner and at an appropriate time, generate understanding. We will argue that various philosophical accounts of mathematical explanation—including those offered by Steiner (1978), Kitcher (1981), and Lange (2014)—are all derivable consequences of a combination of Wilkenfeld's functional account and a modern understanding of human cognitive architecture. Consequently, we argue that Wilken-

✉ Matthew Inglis
m.j.inglis@lboro.ac.uk

1 Mathematics Education Centre, Loughborough University, Loughborough LE11 3TU, UK

2 Rutgers University, New Brunswick, USA

Springer


Synthese (2021) 198 (Suppl 26):S6369–S6392
https://doi.org/10.1007/s11229-019-02234-5

Educational Studies in Mathematics (2022) 111:445–467
https://doi.org/10.1007/s10649-022-10164-2

Check for updates

Do mathematicians and undergraduates agree about explanation quality?

Tanya Evans[1] · Juan Pablo Mejía-Ramos[2] · Matthew Inglis[3]

Accepted: 30 May 2022 / Published online: 29 July 2022

Abstract
Offering explanations is a central part of teaching mathematics, and understanding those explanations is a vital activity for learners. Given this, it is natural to ask what makes a good mathematical explanation. This question has received surprisingly little attention in the mathematics education literature, perhaps because the field has no agreed method by which explanation quality can be reliably assessed. In this paper, we explore this issue by asking whether mathematicians and undergraduates agree with each other about explanation quality. A corpus of 10 explanations produced by 10 mathematicians was used. Using a comparative judgement method, we analysed 320 paired comparisons from 16 mathematicians and 320 from 32 undergraduate students. We found that both mathematicians and undergraduates were able to reliably assess the quality of a set of mathematical explanations. Furthermore, the assessments were largely consistent across the two groups. Implications for theories of mathematical explanation are discussed. We conclude by arguing that comparative judgement is a promising technique for exploring explanation quality.

Keywords Mathematical explanation · Explanation quality · Mathematical practices · Undergraduate mathematics · Comparative judgement

✉ Tanya Evans
t.evans@auckland.ac.nz

Juan Pablo Mejía-Ramos
jpmejia@math.rutgers.edu

Matthew Inglis
M.J.Inglis@lboro.ac.uk

1 Department of Mathematics, University of Auckland, 38 Princes Street, 1010 Auckland, New Zealand

2 Graduate School of Education, Department of Mathematics, Rutgers University, 10 Seminary Place, 08901 New Brunswick, New Jersey, USA

3 Department of Mathematics Education, Loughborough University, Epinal Way, LE11 3TU Loughborough, UK


PLOS ONE
RESEARCH ARTICLE
Stereotype threat, gender and mathematics attainment: A conceptual replication of Stricker & Ward
Matthew Inglis[1], Steven O'Hagan[2*]
1 Centre for Mathematical Cognition, Loughborough University, Loughborough, United Kingdom, 2 School of Mathematics, University of Edinburgh, Edinburgh, United Kingdom
* s.ohagan@ed.ac.uk
Check for updates
OPEN ACCESS
Citation: Inglis M, O'Hagan S (2022) Stereotype threat, gender and mathematics attainment: A conceptual replication of Stricker & Ward. PLoS ONE 17(5): e0267699. https://doi.org/10.1371/journal.pone.0267699
Editor: Jelte M. Wicherts, Tilburg University, NETHERLANDS
Received: June 7, 2021
Accepted: April 14, 2022
Published: May 27, 2022
Copyright: © 2022 Inglis, O'Hagan. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Data Availability Statement: The raw data and analyses scripts are available at https://doi.org/10.17605/OSF.IO/UMJ4H.
Funding: The authors received no specific funding for this work.
Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: At the time the study was conducted, Steven O'Hagan was employed as the Deputy Director of the UK Mathematics Trust. This does not alter our
Abstract
Stereotype threat has been proposed as one cause of gender differences in post-compulsory mathematics participation. Danaher and Crandall argued, based on a study conducted by Stricker and Ward, that enquiring about a student's gender after they had finished a test, rather than before, would reduce stereotype threat and therefore increase the attainment of women students. Making such a change, they argued, could lead to nearly 5000 more women receiving AP Calculus AB credit per year. We conducted a preregistered conceptual replication of Stricker and Ward's study in the context of the UK Mathematics Trust's Junior Mathematical Challenge, finding no evidence of this stereotype threat effect. We conclude that the 'silver bullet' intervention of relocating demographic questions on test answer sheets is unlikely to provide an effective solution to systemic gender inequalities in mathematics education.
Introduction
Mathematics education researchers have long been concerned that mathematics is experienced differently by men and women [1]. This concern is, in part, fueled by gender differences in post-compulsory participation rates in mathematical study and STEM careers [2].
One mechanism which some believe contributes to these observed gender differences in participation is stereotype threat. This account suggests that members of negatively stereotyped groups underperform when that stereotype is salient, perhaps because stereotype-related thoughts place an extra burden on stereotyped individuals' cognitive resources [3]. For example, Steele and Aronson [4] found that black participants underperformed on laboratory tests of verbal ability compared to white participants, but only when reminded of negative stereotypes concerning race and intelligence. Similarly, Spencer, Steele and Quinn [5] found that women performed worse on a laboratory mathematics test than men, but only when they were told that the test usually revealed gender differences in achievement. Subsequently many similar lab-based studies have been conducted: a meta-analysis of 47 such studies showed that women, on average, underperform on laboratory mathematics tests by 0.22 standard deviations when under stereotype threat conditions [6].
PLOS ONE | https://doi.org/10.1371/journal.pone.0267699 May 27, 2022
1 / 12
Educational Studies in Mathem
https://doi.org/10.1007/s1064
Do mathematician
about explanation
Tanya Evans[1] · Juan P
Accepted: 30 May 2022 / Publi
© The Author(s) 2022
Tanya Evans
t.evans@auckland.ac.nz
Juan Pablo Mejía-Ramos
jpmejia@math.rutgers.edu
Matthew Inglis
M.J.Inglis@lboro.ac.uk
1 Department of Mathematics, University of Auckland, 38 Princes Street, 1010 Auckland, New Zealand
2 Graduate School of Education, Department of Mathematics, Rutgers University, 10 Seminary Place, 08901 New Brunswick, New Jersey, USA
3 Department of Mathematics Education, Loughborough University, Epinal Way, LE11 3TU Loughborough, UK
Springer
Synthese (2021) 198 (Suppl 26):S6369–S6392
https://doi.org/10.1007/s11229-019-02234-5
S.I. : ENABLING M
Functional exp
Matthew Inglis[1]
Received: 31 May 2018 / A
© The Author(s) 2019
Research Paper
PAIN
People in pain
Nina Attridge[a,*], Jayne Pickerin
1662 N. Attridge et al. • 160 (2019)
Beau
M

87–109

10.1093/ph

Philosophia

Advance Ac

Research Paper

PAIN

People in pain

Synthese (2021) 198 (Suppl 26):S6369–S6392

https://doi.org/10.1007/s11229-019-02234-5

Educational Studies in Mathem

PLOS ONE

FEATURE ARTICLES

Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned?

Hugues Lortie-Forgues[1] and Matthew Inglis[2]

There are a growing number of large-scale educational randomized controlled trials (RCTs). Considering their expense, it is important to reflect on the effectiveness of this approach. We assessed the magnitude and precision of effects found in those large-scale RCTs commissioned by the UK-based Education Endowment Foundation and the U.S.-based National Center for Educational Evaluation and Regional Assistance, which evaluated interventions aimed at improving academic achievement in K–12 (141 RCTs; 1,222,024 students). The mean effect size was 0.06 standard deviations. These sat within relatively large confidence intervals (mean width = 0.30 *SDs*), which meant that the results were often uninformative (the median Bayes factor was 0.56). We argue that our field needs, as a priority, to understand why educational RCTs often find small and uninformative effects.

Keywords: educational policy; evaluation; meta-analysis; program evaluation

Large-scale randomized controlled trials (RCTs) are now regularly used to evaluate educational interventions. For example, the U.S.-based National Center for Educational Evaluation and Regional Assistance (NCEE) started funding large-scale RCTs in 2002, and the UK-based Education Endowment Foundation (EEF) has funded more than 160 since 2012. This trend is not limited to these two countries: In recent years, funding organizations in the European Union (e.g., European Schoolnet), Japan (e.g., Nippon Foundation), Australia (e.g., Social Ventures), Switzerland (e.g., Jacob's Foundation), Brazil (e.g., Lemann Foundation), and Bangladesh (e.g., BRAC) have also prioritized RCTs in education.

Evaluating the efficacy of educational programs before implementation is important to avoid wasting resources. In medicine, there are many instances where RCTs have shown that promising treatments were ineffective or harmful (Sibbald & Roland, 1998). However, conducting large-scale RCTs is expensive. For example, the EEF spends around £500,000 per trial (EEF, 2015a). Given the growing number of large-scale RCTs in education and their expense, it is important to reflect on how informative this new research focus has been. To our knowledge, no study has systematically evaluated this recent trend. In this article, we use empirical data from two prominent educational funding bodies to evaluate the typical effects produced by large-scale educational RCTs. Our aim is to provide an empirical basis for discussions of the field's efforts to build rigorous scientific evidence.

Randomized Control Trials

RCTs are widely regarded as the "gold standard" for measuring the efficacy of interventions (Pocock, 1983). In their simplest form, participants are randomly assigned to an experimental group that receives the intervention or a control group that receives an alternative treatment or possibly no treatment. The effectiveness of the intervention is then determined by comparing the outcomes between groups. RCTs are highly regarded because compared with other types of studies (e.g., case studies), they ensure that the groups are probabilistically identical at the outset and that any difference in outcome are therefore *caused* by the intervention (assuming that the probability of the difference occurring by chance is sufficiently low).

Unfortunately, not all RCTs are of the same quality (e.g., Higgins et al., 2011). The conclusions of an RCT can be distorted or of limited use if, for example, the sample is too small or not representative, the allocation of the participants is compromised, the outcomes are selectively reported, attrition is ignored, or the outcome measure provides an unfair advantage to the intervention group (e.g., by including material that is taught to the intervention group but not the control group).

[1]University of York, UK

[2]Loughborough University, UK

Educational Researcher, Vol. 48 No. 3, pp. 158–166

DOI: 10.3102/0013189X19832850

Article reuse guidelines: sagepub.com/journals-permissions

[1] Department of Mathematics, University of Auckland, 38 Princes Street, 1010 Auckland, New Zealand

[2] Graduate School of Education, Department of Mathematics, Rutgers University, 10 Seminary Place, 08901 New Brunswick, New Jersey, USA

[3] Department of Mathematics Education, Loughborough University, Epinal Way, LE11 3TU Loughborough, UK

Springer

FEATURE ARTICLES

How Should Educational Effects Be Communicated to Teachers?

Hugues Lortie-Forgues[1], Ut Na Sio[2], and Matthew Inglis[1]

Research findings regarding the effects of educational interventions—typically reported in units of standard deviations (e.g., Cohen's *d*)—are often translated into more intuitive metrics before being communicated to teachers. However, there is no consensus about the most suitable metric, and no study has systematically examined how teachers respond to the different options. We conducted two preregistered studies addressing this issue. We found that teachers have strong preferences concerning effect size metrics in terms of informativeness, understandability, and helpfulness. These preferences challenge current research reporting recommendations. Most importantly, we found that different metrics induce different perceptions of an intervention's effectiveness—a situation that could cause teachers to have unrealistic expectations about what a given intervention may achieve. Implications for how educational effects should be communicated are discussed.

Keywords: communication; decision making; effect sizes; experimental design; teacher knowledge; teachers

It is important to communicate the findings of education research to teachers. One approach is that adopted by the Education Endowment Foundation (EEF) in the United Kingdom and Institute of Education Sciences in the United States. These bodies have commissioned hundreds of randomized control trials (RCTs) and systematic reviews on the most effective teaching practices, which have then been summarized on accessible platforms—the Teaching and Learning Toolkit and What Works Clearinghouse, respectively. Such initiatives have proved very popular: Up to two thirds of schools in England report consulting the Teaching and Learning Toolkit to inform their practice (EEF, 2017), and the What Works Clearinghouse website attracts around 35,000 new users every month (Institute of Education Sciences, personal communication, June 1, 2020). Not surprisingly, similar initiatives have been appearing around the world, such as Evidence for Learning (Australia) and SUMMA (Latin America and the Caribbean). Despite these efforts to improve the availability of evidence in education, little research has examined how to present education research findings in ways that maximize the ability of teachers to make informed decisions. This omission is surprising given the large number of teachers who engage with education research (e.g., Barton & Tindle, 2019), given that substantial efforts have been made to increase teacher's use of research in their practice (e.g., Farley-Ripple et al., 2018; Goldacre, 2013), and given that many initiatives that summarize the impact of educational interventions consider teachers to be one of their target audience (e.g., EEF, 2018; Evidence for Learning, n.d.; SUMMA, n.d.).

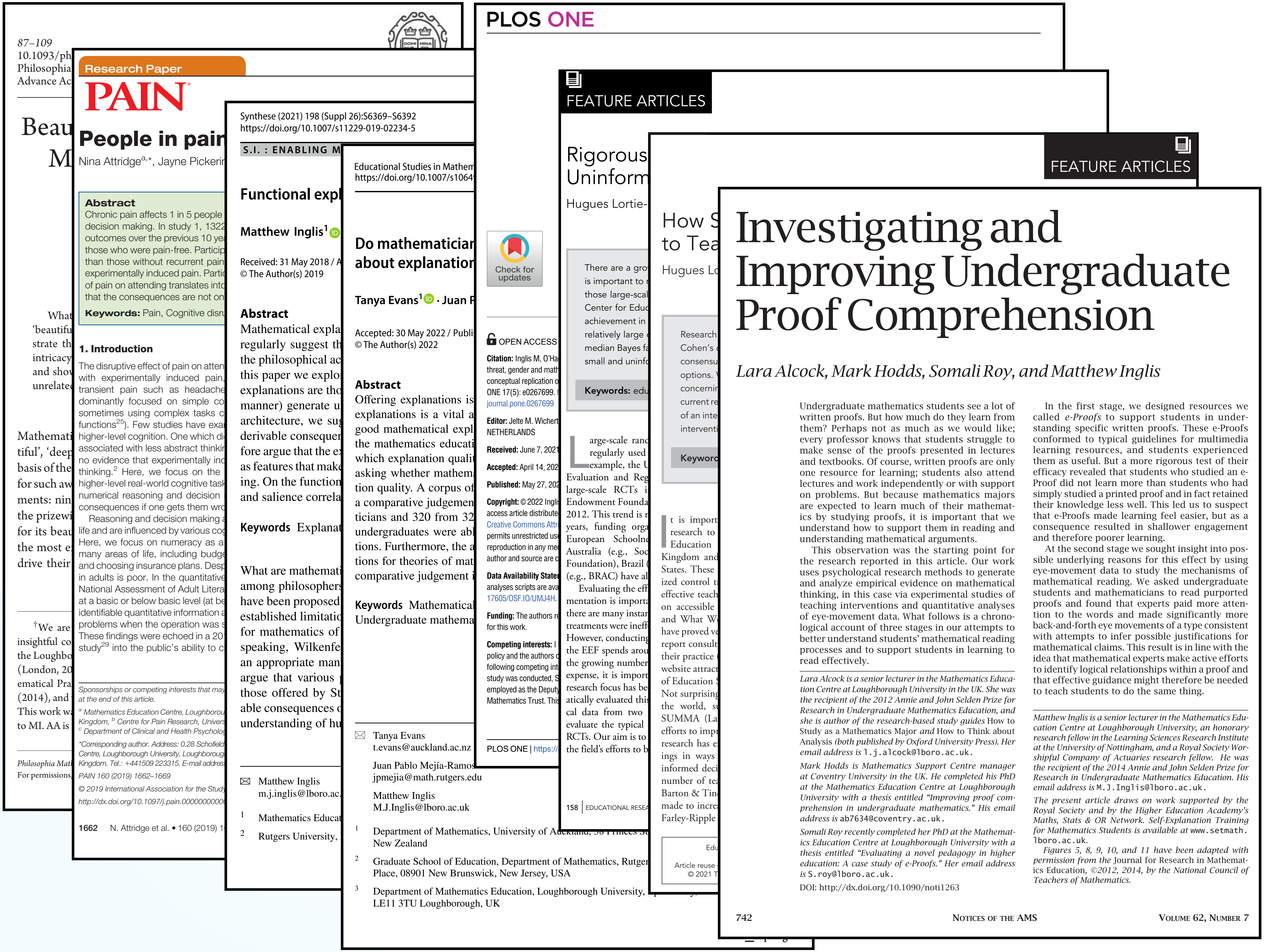
Translation of Effect Sizes in Education

In education research, an intervention's impact is typically reported in units of standard deviations (e.g., Cohen's *d*; see Kraft, 2020, for an overview of effect sizes of educational interventions in relation to their cost and scalability). Because this measure is hard to interpret, it is generally translated into a more relatable metric before being reported to practitioners. Many alternative metrics have been proposed (e.g., Lipsey et al., 2012), but to date, there is no consensus about the metric best suited for communication with practitioners. For example, the Teaching and Learning Toolkit, Evidence for Learning, and SUMMA translate effects into additional student months of progress, while the What Works Clearinghouse reports effects as percentile gains (referred to as the improvement index).

Researchers have argued that some metrics are better than others. In their effect size interpretation guidelines, Valentine and Cooper (2003) recommended reporting raw mean

[1]Loughborough University, Loughborough, UK
[2]University of Sheffield, Sheffield, UK

Educational Researcher, Vol. 50 No. 6, pp. 345–354
DOI: 10.3102/0013189X20987856
Article reuse guidelines: sagepub.com/journals-permissions



Investigating and Improving Undergraduate Proof Comprehension

Lara Alcock, Mark Hodds, Somali Roy, and Matthew Inglis

Undergraduate mathematics students see a lot of written proofs. But how much do they learn from them? Perhaps not as much as we would like; every professor knows that students struggle to make sense of the proofs presented in lectures and textbooks. Of course, written proofs are only one resource for learning; students also attend lectures and work independently or with support on problems. But because mathematics majors are expected to learn much of their mathematics by studying proofs, it is important that we understand how to support them in reading and understanding mathematical arguments.

This observation was the starting point for the research reported in this article. Our work uses psychological research methods to generate and analyze empirical evidence on mathematical thinking, in this case via experimental studies of teaching interventions and quantitative analyses of eye-movement data. What follows is a chronological account of three stages in our attempts to better understand students' mathematical reading processes and to support students in learning to read effectively.

Lara Alcock is a senior lecturer in the Mathematics Education Centre at Loughborough University in the UK. She was the recipient of the 2012 Annie and John Selden Prize for Research in Undergraduate Mathematics Education, and she is author of the research-based study guides How to Study as a Mathematics Major *and* How to Think about Analysis *(both published by Oxford University Press). Her email address is* l.j.alcock@lboro.ac.uk.

Mark Hodds is Mathematics Support Centre manager at Coventry University in the UK. He completed his PhD at the Mathematics Education Centre at Loughborough University with a thesis entitled "Improving proof comprehension in undergraduate mathematics." His email address is ab7634@coventry.ac.uk.

Somali Roy recently completed her PhD at the Mathematics Education Centre at Loughborough University with a thesis entitled "Evaluating a novel pedagogy in higher education: A case study of e-Proofs." Her email address is S.roy@lboro.ac.uk.

DOI: http://dx.doi.org/10.1090/noti1263

In the first stage, we designed resources we called *e-Proofs* to support students in understanding specific written proofs. These e-Proofs conformed to typical guidelines for multimedia learning resources, and students experienced them as useful. But a more rigorous test of their efficacy revealed that students who studied an e-Proof did not learn more than students who had simply studied a printed proof and in fact retained their knowledge less well. This led us to suspect that e-Proofs made learning feel easier, but as a consequence resulted in shallower engagement and therefore poorer learning.

At the second stage we sought insight into possible underlying reasons for this effect by using eye-movement data to study the mechanisms of mathematical reading. We asked undergraduate students and mathematicians to read purported proofs and found that experts paid more attention to the words and made significantly more back-and-forth eye movements of a type consistent with attempts to infer possible justifications for mathematical claims. This result is in line with the idea that mathematical experts make active efforts to identify logical relationships within a proof and that effective guidance might therefore be needed to teach students to do the same thing.

Matthew Inglis is a senior lecturer in the Mathematics Education Centre at Loughborough University, an honorary research fellow in the Learning Sciences Research Institute at the University of Nottingham, and a Royal Society Worshipful Company of Actuaries research fellow. He was the recipient of the 2014 Annie and John Selden Prize for Research in Undergraduate Mathematics Education. His email address is M.J.Inglis@lboro.ac.uk.

The present article draws on work supported by the Royal Society and by the Higher Education Academy's Maths, Stats & OR Network. Self-Explanation Training for Mathematics Students is available at www.setmath.lboro.ac.uk.

Figures 5, 8, 9, 10, and 11 have been adapted with permission from the Journal for Research in Mathematics Education, ©2012, 2014, by the National Council of Teachers of Mathematics.

Thanks!

Web: http://mcg.lboro.ac.uk/mji/

Email: m.j.inglis@lboro.ac.uk

Twitter: @mjinglis