Bentley & Yates, *Cogent Education* (2017), 4: 1297213
https://doi.org/10.1080/2331186X.2017.1297213

cogent education

CURRICULUM & TEACHING STUDIES | RESEARCH ARTICLE

Facilitating proportional reasoning through worked examples: Two classroom-based experiments

Brendan Bentley[1]* and Gregory C.R. Yates[1]

Received: 21 August 2016
Accepted: 15 February 2017
Published: 14 March 2017

*Corresponding author: Brendan Bentley, School of Education, University of South Australia, Adelaide, Australia
E-mail: brendan.bentley@unisa.edu.au

Reviewing editor:
John Lee, Hong Kong Institute of Education, Hong Kong

Additional information is available at the end of the article

Abstract: Within mathematics teaching, ways to help students resolve proportional reasoning problems remains a topical issue. This study sought to investigate how a simple innovative procedure could be introduced to enhance skill acquisition. In two classroom-based experiments, 12-year-old students were asked to solve proportional reasoning mathematics problems, on four occasions, over a two-week period. On the second occasion, students worked either with or without the benefit of worked examples. The examples demonstrated a unitising strategy in the context of solving proportional reasoning missing value problems. Students exposed to the worked examples improved scores on subsequent tests. The worked example instruction was (a) mediated entirely through booklets, (b) effective with both low- and high-SES students and (c) represents a promising approach to teaching within an area that habitually presents many challenges for the general classroom teacher.

Subjects: Mathematics Education; Cognitive Psychology; Thinking, Reasoning & Problem Solving

Keywords: proportional reasoning; mathematics education; cognitive load theory; self-efficacy

ABOUT THE AUTHORS

Brendan Bentley specialises in teaching Science and Mathematics Education at the University of South Australia. He has been a school leader and educator in excess of 30 years. His general interests are in cognition, cognitive load theory and learning. He has published in area of civics and values education, and is an active researcher in the curriculum disciplines of science and mathematics education in particular proportional reasoning.

Gregory C.R. Yates is an adjunct senior lecturer at the University of South Australia. He has research publications in the area of cognitive psychology and social learning. He is the co-author with John Hattie of *Visible learning and science of how we learn (Routledge 2013).*

PUBLIC INTEREST STATEMENT

It has been found that many individuals, adults and children alike, experience difficulty when they need to think and reason in terms of mathematical proportions. This trait is called proportional reasoning in traditional curriculum statements, and can be indexed by the question, "A patient's dosage is 75 mm per week. The drug is diluted 1 to 50 with water. How much water is needed to mix a month's supply?" Many commentators within this area note the need to help teachers identify effective teaching methods. The present article argues that such teachers can subtly introduce their students to the heuristic strategy of unitising. This is a breakdown method that can be applied to many types of problems involving proportional reasoning. We report on two classroom-based experiments where 12-year-old students were found to score more highly on tests of proportional reasoning after being encouraged to analyse worked examples on paper which showed how to apply the unitising method in solving proportional reasoning problems.

© 2017 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

1. Introduction

It costs $3.25 for 90 g of Arid deodorant, or $4.35 for 140 g. *Which option is the better buy*? Apparently, the majority of the adult population appear unable to solve similar problems correctly, even when supplied with paper and pencil (Capon & Kuhn, 1982). In this paper, we advance the idea, that, in accord with the worked example principle, students at the primary/lower secondary level will benefit from exposure to such problems that have already been solved using a clear unitising strategy. To assay this notion, we conducted two experiments: One in a high socio-economic status (SES) school, and the other in a low-SES school location. Worked examples were used to demonstrate the strategy of unitising, i.e. to help students adopt a step-by-step analysis through which an easily recognisable unit can firstly be articulated, and then utilised to solve the problem. We interpret the findings through the lens of cognitive load theory. In this paper, we (a) briefly review the theory of worked examples, (b) indicate how using worked examples can convey an effective problem solving strategy and (c) describe the findings of the two classroom-based experiments.

1.1. Impact of studying worked examples

A considerable body of knowledge supports the notion that, when learners are relative novices, they benefit enormously from exposure to modelled examples (Renkl, 2005; Sweller, 1988; van Gog & Rummel, 2010). Examples provide the means through which a learner can acquire new skills, and then begin to apply these skills to problems they have not previously encountered. Studies have established that, through exposure to clear worked examples, novice learners will learn more quickly, and report lower levels of task difficulty (or cognitive load), relative to similar peers who were not exposed to worked examples. The reduction in load is, in turn, associated with enhanced ability to attend to and retain crucial information, and then to apply this knowledge in the service of problem solving. Indeed, the term 'worked example effect' is used to describe contexts in which learners are relatively advantaged by studying examples, but relatively disadvantaged by being expected to learn through solving problems without exposure to exemplars (Sweller, van Merrienboer, & Paas, 1998; van Loon-Hillen, van Gog, & Brand-Gruwel, 2012).

The research literature on the worked example effect has become voluminous and subject of several reviews (Hattie, 2009; Renkl, 2014). The majority of such studies have been conducted with adolescents and young adults—for instance with tertiary students, individuals undertaking vocational training programs, and high-school students. However, we were able to locate only five studies demonstrating the impact of worked examples in the context of primary/lower secondary school teaching (Jitendra, Star, Dupuis, & Rodriguez, 2013; Mwangi & Sweller, 1998; Retnowati, Ayres, & Sweller, 2010; Star & Rittle-Johnson, 2009; van Loon-Hillen et al., 2012).

Retnowati et al. (2010) used worked examples written in booklets to teach 12-year-old students how to apply geometric theorem problems in a mathematics classroom context. Following on from a formal lesson presentation, students worked either individually or in small groups, and this variable was of little importance. However, students were exposed either to worked examples presented in the training booklet or worked on problems without exposure to such examples. The worked example group exceeded the non-exposure group on both acquisition and transfer tests, with effect sizes of 1.28 and .89 respectively (Cohen's *d*). These students also indicated that being exposed to worked examples assisted them to understand the material taught. A feature of the training booklets used was that students were given problem pairs (i.e. for each problem pair, the students studied the worked example, before solving a similar problem by themselves).

Star and Rittle-Johnson (2009) established that 10 and 11-year-old students could considerably improve their mathematics estimation skills through studying different types of worked examples (i.e. samples of how others made estimations). They found one particularly effective method was to enable students to compare alternative estimation methods physically side by side. However, the study did not involve a control group design and it was unclear as to the extent of the effectiveness of the worked examples.

Mwangi and Sweller (1998) used worked examples to successfully teach nine-year-old students to solve two-stage arithmetic word problems. Although the worked example group and a control group achieved similar levels of mastery, it was found that exposure to worked examples enabled students to learn both more quickly and with fewer errors (d = .9). The impact was found to be strongest when examples were presented in a well-integrated format, as distinct from a format that involved the students' attention being divided between different sources of information on the page (d = 1.3).

Van Loon-Hillen et al. (2012) reported on a school-based project with nine-year-olds in which enrichment materials featuring worked examples were associated with a mathematics curriculum unit in one classroom, but not in a comparison classroom. Although the two classes did not differ on achievement test outcomes, students in the worked example enrichment class required less time to achieve mastery (29 vs 34 min on average, d = .8), with 80% of the students indicating that wanted to continue learning through using such examples.

Furthermore, Jitendra et al. (2013) reported on a 6-week intervention with classes of 13-year-old students markedly improved their proportional reasoning scores, relative to other classes that did not participate (d = 1.27). The intervention focussed upon identifying the schematic structure underpinning proportional reasoning problems, analysing visual representations of worked examples, and applying procedural heuristics.

Hence, although there exists general agreement that students will benefit from studying worked examples in mathematics, such a notion is not supported by a substantive database with respect to primary age students. Furthermore, we could not locate any controlled studies showing that exposure to worked examples alone would enhance proportional reasoning capabilities in primary/lower secondary age students.

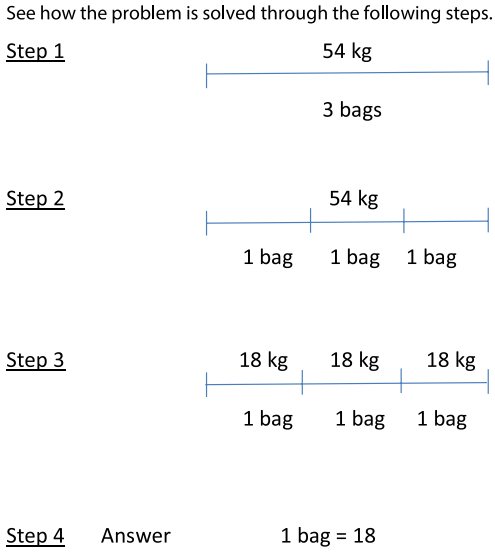
1.2. Teaching proportional reasoning skills at the middle school level

Within a school context, students use proportional reasoning when they engage in rational number which includes, fractions, percentages, decimals and other mathematic areas such as trigonometry, cooking, scale drawing and as well as other academic disciplines (Lamon, 2007; Norton, 2005). Lesh, Post, and Behr (1988) have highlighted several conceptual 'trouble spots' (p. 95), these are equivalent fractions, long division, place value and percentages, measurement and conversions, and ratios and rates. Of interest, these conceptual trouble spots are commonly experienced within the middle school curriculum.

Despite proportional reasoning being seen as an availing skill, fundamental to many areas of human endeavour, two glaring issues stand out whenever educators begin to reflect upon common practices and general pedagogic approaches, employed to convey knowledge in this area. These issues are (a) students experience difficulty in achieving mastery in this area, notably when expected to shift from additive to multiplicative reasoning, and (b) the general classroom teacher finds this curriculum area to pose many problems that render teaching challenging (Sowder, 2007). Significantly, multiplicative reasoning has been identified as an essential component of proportional reasoning (Lamon, 2006).

Within the Australian context, teaching young students to reason proportionally has been viewed as a problematic area of the mathematics curriculum (Dole, 2008; Hilton, Hilton, Dole, & Goos, 2016). This matter is compounded, in that the majority of students are instructed in their early mathematics skills by non-specialist teachers with limited access to specialist advisors in this area. Currently, the recognised methods used to teach proportional reasoning appear largely constrained to a range of algorithms and mechanical procedures (Lamon, 2006, 2007), coupled with generalised encouragement. More effective ways to teach proportional reasoning strategies need to be investigated, particularly with a view to assisting non-specialist teachers to develop effective methodologies.

Figure 1. Worked example demonstrating unitisation strategy used in treatment for WE group.



In general, there have been admonitions about the need to encourage students to "*think flexibly*", and duly "*solve the problem through carefully thinking about it*". Whilst it is customary for students to receive much encouragement about how to solve difficult mathematics problems, and for teachers to view such problems as natural extension items for individual work, nevertheless, it is possible to express the view that within the average Australian classroom very little direct teaching is likely to take place concerning reasoning in proportional thinking (Hilton et al., 2016).

As an experienced (30 years) teacher at the primary/lower secondary school level, the first author is accustomed to teaching proportional reasoning to primary/lower secondary school students through helping them to develop unitising strategies. In this connection, unitising means to reduce the given information to easily recognisable units, often to a unit of one. We note one way to 'solve' the Arid deodorant problem (cited in the Introduction), within an Australian shop is to read the shelf label carefully, where the unitised price per gram is displayed, albeit in small font.

It is known that young children exhibit elementary notions of proportionality (such as dividing a pizza into equal shares, or perceiving cartoons as being drawn out of proportion). Further, developmental research has disclosed that young children may possess an intuitive grasp of unitisation (Lamon, 2007). However, for many young students developing this as a first-line stratagem within a "school-like" situation represents a significant conceptual advance (English & Mulligan, 2013). Once mastered, unitisation strategies can be readily applied to qualities such as fractions, rates and ratios. In addition, unitisation can become a foundation point for more complex strategies such as build-up, unit rate or factor of change (Artut & Pelen, 2015; Cramer, Bezuk, & Behr, 1989).

By its very nature, learning to unitise implies carefully dissecting given examples. It is through analysis of well-articulated examples that steps can be identified which describe the application of the strategy. These steps are shown in Figure 1, which, in effect, becomes a worked example. It has been established, through cognitive load research, that the most effective worked examples are ones that highlight critical cues, are relatively simple to follow, do not present elaborated information and which encourage a level of immediate participation (Renkl, 2014).

The model as depicted in Figure 1 was developed by the first author through a series of trials with young students before being used in the present studies. Underpinning this approach are two notions. The first is that proportional reasoning problems possess an inherent level of difficulty since they involve simultaneously working with multiple elements (such as different variables e.g. time

and distance). Secondly, it is assumed that cueing a simple procedural strategy will enhance student performance and reduce task difficulty level.

1.3. The present studies

We hypothesised that students who are exposed to worked examples depicting the strategy of unitising, would show a significant increase in the ability to solve proportional reasoning problems. Although students at this level ostensibly have been taught proportional reasoning, we sought to investigate (a) how a simple procedure could be introduced into practice opportunities on one day, and (b) if this procedure would affect students on following test sessions.

To investigate, if exposure to worked examples would enhance proportional reasoning skills, a training booklet was devised and then used in two classroom-based experiments. The booklet featured 10 such fully worked examples, where a student first studied an example, and then attempted to replicate the step-by-step procedure on the adjacent page with a parallel problem. It would require about 45 min to work through the booklet, which involved 10 worked examples interleaved with 10 test items.

Both experiments used identical methodological procedures. It was arranged that one group of students would be exposed to the training booklet, whereas a comparable group would work on the identical test problems but without exposure to the examples. In addition, we tapped (a) student confidence (self-efficacy, prior to attempting items), and (b) their experienced cognitive load, by asking students to rate the task difficulty of each item. We predicted that students exposed to worked examples would outperform the other students on measures of proportional reasoning, on both immediate tests, and one-week delayed tests. We predicted that task difficulty could be reduced through studying examples, but no predictions were possible in connection with confidence self-efficacy ratings.

2. Experiment 1

2.1. Method

2.1.1. Participants

In all, 54 students participated in the first study (E1). All attended a year seven class in a government school located in a relatively affluent suburb in Adelaide. However, data from eight students who obtained ceiling scores on the pretest measure, and two with diagnosed special needs, were not used in analyses. The 44 remaining students ranged in age from 11.0 to 12.5, with a median of 12.0 years. One student was absent on the final testing day, T4. The 54 participants stemmed from a sample of approximately 80 students from three Year 7 classes at the school on the basis of ethics clearance. The three classrooms were adjacent, and classes were run concurrently during mathematics instructional time, using a collaborative teaching model. During the study, the students were distributed across the three classrooms per normal teaching arrangements. For the purposes of the project, students who did not participate were allocated to one classroom. Students who did participate were randomly allocated into either the Worked Example (WE, $n = 21$, 11 boys, 10 girls) or the Control (C, $n = 23$, 10 boys, 13 girls) groups, as assigned to different classrooms. These arrangements were facilitated by the teachers as part of their normal teaching routines in commencing mathematics classes.

2.1.2. Materials

The procedures were administered through prepared booklets. The cover page of each booklet was blank except for a space to record the participant's age, name, and gender. The pretest booklet used at T1 was identical to that used at T4. It consisted of 10 proportional reasoning problems, or items, presented 1 per A4 page. Five items concerned missing values (MV), and five concerned comparison problems (CP). A sample missing value item is *"Sally rides her bike 12 km from school to home in 2 hours. How far will she travel in 1 hours at the same speed?"* A sample comparison problem is "*Sam*

Table 1. Experimental sequence

Group	T1	T2	T3	T4
	Baseline, day 1	**Treatment, day 7**	**Posttest, day 8**	**Posttest, day 15**
Worked examples (WE)	Pretest (5 MV, 5 CP)	10 MV problems linked to exposure to 10 isomorphic worked examples	Posttest (5 CP, 5 MV)	Same booklet as pretest
Control (C)	Pretest (5 MV, 5 CP)	10 MV no worked example exposure	Postest (5 CP, 5 MV)	Same booklet as pretest

Note: MV means missing value problem and CP means comparison problems.

Figure 2. Cognitive load self-assessment measure.

Too easy	Very very easy	Very easy	Easy	Neither difficult nor easy	Difficult	Very, difficult	Very, very difficult	Too difficult

rides his skateboard 6 km from school to home in 3 hours. Lily takes 4 hours to ride her skateboard 4 km? Who is the fastest skateboarder?"

The booklet at T2 (i.e. the treatment manipulation) was designed to have items parallel to the booklet used at T3 but used variations in the values used within each item, in an attempt to make the test slightly easier. The same booklets were used with both C and WE groups at T1, T3 and T4. However, the C and WE group had separate booklets at T2 (see Table 1). It should be noted that although the WE treatment booklet featured 10 missing value items, the booklets used at T1, T3 and T4, each used five missing value items and five comparison problems.

The specific items used in all booklets were devised in consultation with the teachers at the school. This ensured that (a) the items were consistent with curriculum content, as had been covered in earlier lessons, and (b) the control group experience reflected existing practices already in place within the school.

Within the test booklets for T1, T3 and T4, the students were asked to respond to two additional questions per item, on each page. These were intended as measures of (a) self-efficacy, and (b) cognitive load. Upon reading each item, participants were requested to judge self-efficacy by answering the question: "*Do you think you can solve this problem?*" using a five-point scale: "*no/probably not/cannot tell/probably can/yes*". For analysis, responses were summed across all 10 items to score self-efficacy out of 50 on each occasion.

At the bottom of every page, as a measure of load, participants were requested to respond to the question, "*How difficult did you find this problem?*" using a nine-point scale anchored at "*too easy*" and "*too difficult*". Figure 2 shows the actual scale as it appeared in the booklets. Summing across 10 items enabled load to be scored out of 90 on each occasion. This measure, as suitable for young students, was based on scales as used in cognitive load studies, such as Ayres (2013). In comparing this specific measure to other possible scales used to monitor cognitive load, Ayres specifically noted: "the difficulty scale is particularly sensitive to fluctuations in intrinsic cognitive load" (p. 119).

The booklets used in the treatment at T2 employed only missing value problems. Ten such items were presented as a sequence which progressed from easy to difficult. The C group students were presented with one item per A4 page. The assessment question about cognitive load was at the bottom of each page. (Note the self-efficacy item was not used at T2).

In the case of the WE group, the students were asked to complete the identical ten problems as used in the C group booklet. However, for each item, a variant of the problem was shown on the immediately preceding page. This page displayed the unitising strategy, that is, four specific steps to employ when calculating missing value proportional reasoning problems, as well as the specific solution. The steps were shown as exemplars. The students were expected to complete the parallel (or isomorphic) item on the next page, by using the same strategy. Thus, over the course of the session, the WE students were exposed to 10 such worked examples and asked to replicate the unitisation procedure as workings on the facing page. In contrast, the C students worked on the same problems, without exposure to the isomorphically paired examples.

2.1.3. Procedure

Through existing teaching arrangements at the school, it was possible to move students across classrooms with relative ease. Once approval procedures were completed, participants were randomly allocated into groups prior to pretest. Consistent with normal teaching procedures, the students were distributed across the three teaching spaces, with the worked example, control and non-participant groups in adjacent classrooms. The process was managed through teachers directing participants to their allocated rooms, in accord with routines. At each phase, students were handed individual work booklets shortly after entering the room and being seated.

The three teachers who administered the procedures effectively worked together. The two teachers assisting in the study used a prepared script. The script given at T1 explained that the school was assisting "the university" to investigate children's thinking, and students were being asked to show their work in special booklets. On each testing situation (across the four times) the teachers' scripts encouraged students to read the booklets carefully and to try hard. Students were seated at desks with pencils and instructed to work silently in order to demonstrate their thinking skills. In each room, at all times, two adults (one teacher, one researcher) were present to monitor the class. Each session took the entire lesson period of 50 min, after which booklets were collected and students dismissed.

The procedure was repeated the following week at T2, with individuals directed to either the C or WE classroom. The teachers' actions (and verbal scripts) were ostensibly identical across rooms. Upon entering the room, students were instructed to sit down and begin work on the 'university booklets'. Again, they were reminded to use their thinking skills, to read carefully, and to try hard. However, any discussion was not permitted. Classes were dismissed after 50 min, and the booklets collected.

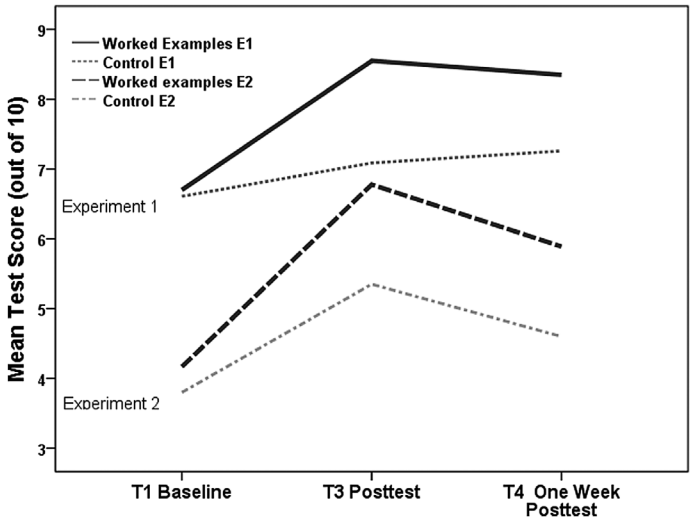
All testing took place in the morning at the school's customary time for mathematical instruction. Procedures were repeated the day following the learning phase, and then again after seven days. These are referred to as T3 and T4, respectively. The procedures ran smoothly, the rooms remained quiet, and students were closely monitored. No procedural or conduct problems were experienced. The students undertook their normal classroom lessons on the non-test days during the experimental period and these lessons did not contain content relating to proportional reasoning.

2.2. Results

2.2.1. Preliminary analyses

Preliminary analyses indicated that, although the baseline test consisted of five MV items and five CP items, principal components analysis revealed these types could not be differentiated meaningfully. With all 10 items contributing, they constituted a coherent scale with an alpha coefficient of .72. At T1, the mean was 6.5, with skewness and kurtosis less than 1.0. Scores at T3 and T4 were found to correlate at .71, which justified the use of repeated measures analyses. Preliminary ANOVAs were computed with gender included as an independent variable. It was apparent that no significant main or interaction effects were associated with this factor. Further, principal component analyses conducted on the self-efficacy and load ratings revealed that, in each case, single-factor resolutions

Figure 3. Mean proportional reasoning test scores in two experiments.



were appropriate, with resultant scales possessing acceptable psychometric properties in terms of alpha, skewness, and kurtosis coefficients.

2.2.2. Findings

A 2 × 3 ANOVA with treatment and time (T1, T3, T4) as factors indicated significant interaction effect, F (2, 82) = 5.4, p = .014, as well as a significant effect for treatment, F (1, 41) = 5.2, p = .028. The interaction indicated the impact of treatment hinged upon test occasion, as is shown in Figure 3, and also Table 2. As shown by means analyses in Table 2, the WE group clearly outperformed the control group on both post-tests, (*p*s of .004 and .026), with effect sizes (Hedges' *g*) of .9 at T3, and .68 at T4. (Note: *g* is akin to Cohen's *d*, but embodies a correction for low numbers).

It was also found (see Table 2) that the WE students reported lower levels of cognitive load, and higher levels of self-efficacy, relative to the C students who did not have access to the worked examples, with all contrasts at T3 and T4 being significant (p < .05, see Table 2, and Figures 4 and 5).

Table 2. Experiment 1 means and deviations of worked example and control groups

Test occasion	Control (*n* = 23)	Worked example (*n* = 21)	*F* ratio	*df*	*p*-value	*g*
T1 total	6.61 (1.5)	6.38 (2.2)	<1	1, 42	.67	.12
T2 total	8.35 (1.9)	9.42 (1.0)	5.16	1, 42	.028*	.68
T3 total	7.10 (1.7)	8.50 (1.3)	9.17	1, 42	.004**	.9
T4 total	7.26 (1.7)	8.35 (1.4)	5.36	1, 41	.026*	.68
T1 load	40.09 (11.6)	39.60 (10.6)	<1	1, 42	.89	.04
T2 load	30.22 (15.8)	29.29 (11.4)	<1	1, 42	.35	.07
T3 load	37.52 (15.9)	28.62 (10.8)	4.65	1, 42	.037*	.66
T4 load	35.48 (12.9)	27.25 (11.2)	4.92	1, 41	.032*	.67
T1 self-eff	38.17 (8.0)	40.70 (5.0)	1.55	1, 42	.22	.37
T3 self-eff	39.65 (8.8)	44.29 (4.5)	4.73	1, 42	.035*	.64
T4 self-eff	38.70 (8.6)	44.10 (5.4)	5.87	1, 41	.02*	.73

Notes: (a) The second and third columns represent means and deviations. (b) Total scores are out of 10, (c) Load scores are out of 90, where high scores reflect high perceived difficulty, (d) Self-efficacy is out of 50 where high scores represent higher confidence, (e) *df* is 41 at T4 as one student was absent, (f) effect sizes expressed as Hedges' *g*, and (g).
*Significance level at p < .05.
**Significance level at p < .01.

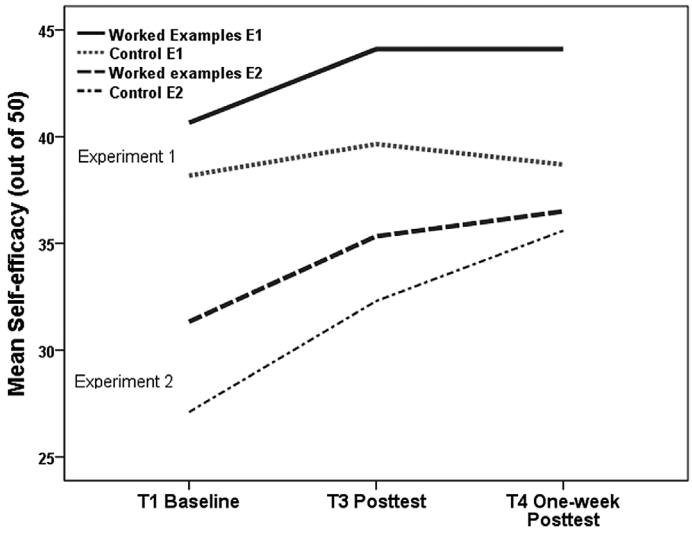


Figure 4. Mean self-efficacy rating in two experiments.

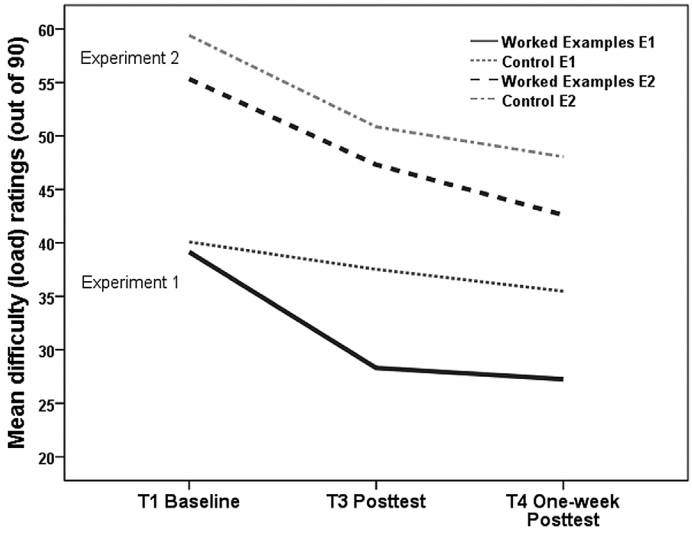


Figure 5. Mean task difficulty (cognitive load) ratings in two experiments.

The mean figures in Table 2 indicate that the items used at T2 were indeed easier than pretest items (since T2 scores were higher than T1 scores even for the C group alone, t (22) = 4.6, $p < .01$). However, it is especially interesting to note that, after spending almost 4 hours working on proportional reasoning problems, the students in the C group failed to significantly improve their performance from T1 to T4 (paired t (22) = 1.8, *ns*), despite the tests using identical items. Similarly, these C group students reported similar levels of self-efficacy on the two occasions ($p = .69$) (see Figure 4), even though they did report lower levels of load from T1 to T4, paired t (22) = 2.9, $p < .01$ (see Figure 5).

3. Experiment 2

In the first experiment, students who studied a booklet depicting 10 worked examples depicting the unitisation strategy demonstrated shifts in their test scores, relative to peers encouraged to try hard on similar items. It was deemed important to attempt a replication since (a) students in Experiment 1 attended a high-SES school, (b) 8 students had obtained ceiling scores on the baseline measure and appeared competent in employing unitising skills already, and (c) scores on the T1 pre-treatment phase were (at a mean of 6.5 out of 10) considerably higher than expected on basis of our pilot work with similar-aged students. Hence, although the paper-based treatment evidently 'worked', it was important to establish its effectiveness did not depend on upon usage with students already advanced in their academic skills.

Table 3. Experiment 2 means and deviations of worked example and control groups

Test occasion	Control (*n* = 20)	Worked example (*n* = 18)	*F* ratio	*df*	*p*-value	*g*
T1 total	3.80 (2.0)	4.17 (2.2)	<1	1, 36	.80	.17
T2 total	4.70 (2.4)	8.44 (2.1)	25.3	1, 36	.001**	1.62
T3 total	5.35 (1.4)	6.78 (1.8)	7.6	1, 36	.009**	.87
T4 total	4.60 (1.7)	6.06 (2.0)	5.41	1, 36	.023*	.77
T1 load	59.40 (16.0)	55.33 (16.0)	<1	1, 36	.37	.25
T2 load	53.00 (22.7)	41.39 (23.5)	2.4	1, 36	.13	.49
T3 load	50.85 (15.5)	47.31 (21.1)	<1	1, 36	.55	.19
T4 load	48.05 (14.3)	42.61 (19.2)	<1	1, 36	.33	.32
T1 self-eff	27.10 (6.5)	31.33 (9.9)	<1	1, 36	.12	.50
T3 self-eff	32.30 (9.3)	35.33 (10)	<1	1, 36	.34	.31
T4 self-eff	35.60 (8.5)	36.50 (9.2)	<1	1, 36	.76	.1

Notes: (a) The second and third columns represent means and deviations. (b) Total scores are out of 10, (c) Load scores are out of 90, where high scores reflect high perceived difficulty, (d) Self-efficacy is out of 50 where high scores represent higher confidence, (e) the final columns represents effect size expressed via Hedges' *g*, and (f).
*Significance level at $p < .05$.
**Significance level at $p < .01$.

3.1. Method

3.1.1. Participants

In all, 38 students participated in all phases of Experiment 2 (E2). They stemmed from three year seven classes in a government school located in suburban Adelaide. They ranged in age from 11.0 to 13.3, with a median of 12.3 years. Eighteen students (7 boys and 11 girls) were assigned randomly to the WE group, and 20 students (9 boys and 11 girls) assigned to the C group. The three classrooms were located within the same classroom complex, and classes ran concurrently during mathematics instructional time, using a collaborative teaching model. During the study, the students were distributed across classrooms consistent with normal teaching arrangements at the school.

On official Australian Government ICSE ratings, the school used in Experiment 1 was in the top 5% of all national schools in terms of socio-economic advantage. The school used in Experiment 2 was ranked in the lowest quartile.

3.1.2. Procedure

The procedure employed in the second experiment replicated that of Experiment 1. The school involved in Experiment 2 had three adjacent classrooms for year seven classes, and students were accustomed to moving between rooms for different purposes. The three teachers espoused a strong dedication in assisting their students in basic skills and mathematics. When examining the experimental tests, they each remarked that, although they were suitable, some specific items would likely be difficult ones for many students.

3.2. Results

Preliminary analyses, using ANOVA, indicated that gender did not significantly impact test scores in terms of either a main or interaction effect. Again, it was apparent that MV items did not differentiate from CP items, and a coherent scale was evident on the initial test occasion with an alpha of .71, and all items contributing, except for Item 9 which was not solved by any student at T1.

The relevant means are shown in Figure 3 and Table 3. Through means testing (see Table 3) it was found that WE students did not differ from C students on the T1 baseline test, but outperformed the C group at both T3 ($g = .87$), and T4 ($g = .77$).

Differences between the C and WE groups on measures of both self-efficacy and load failed to achieve significance. It was apparent that, over time, the students increased in their self-efficacy, and reported lower levels of load. But, these changes were, as expressed through ANOVA coefficients, unrelated to the worked example group treatment factor.

However, it was still important to assay for effects of untutored practice alone, since the two groups diverged markedly on the T2 test. When students in the control group alone (i.e. no strategy training) were examined, it was evident they had improved their reasoning scores from T1 to T4, paired t (19) = 2.27, p = .035. Further, they reported reduced levels of load (paired t (19) = 4.2, p < .01), as well as enhanced levels of self-efficacy, paired t (19) = 4.7, p < .01. Hence, for this group of students, untutored practice on proportional reasoning over four periods was associated with increased test scores, increased self-efficacy, and reduced load. However, achievement increases were considerably higher in those exposed to worked examples. The increment in the C group was from a mean of 3.8 to 4.6, i.e. 21% increase, representing a repeated measures Cohen effect size d = .43. The corresponding change within the WE group were means of 4.2 and 6.1 (45% increase), representing an effect size d = .9.

3.3. Comparing findings from both experiments

3.3.1. Achievement outcomes

Direct comparisons across the two studies are possible since the tests and designs were parallel and student ages were comparable. The major difference being the socio-economic status (SES) of the school districts; one being clearly economic advantaged (high-SES); the other being less advantaged (low-SES). Site location (E1 vs E2) accounted for huge effects at T1 in terms of achievement scores, F (1, 80) = 49, p < .01, d = 1.54. Similar magnitude effects were evident for self-efficacy (d = 1.34), and for cognitive load (d = 1.44).

In both experiments, students who were exposed to worked examples showed increases in their proportional reasoning scores. Through worked examples, the E1 high-SES group increased their mean scores on a 10-point test from around 6.4 to 8.4. The E2 low-SES group increased from around 4.2 to 6.1.

It is apparent that at the T3 point, the WE group from the low-SES school scored at the same level as the C group from the high-SES school (means contrast via LSD procedure, p = .53). By T4 the picture was unclear as the low-SES group scored lower than the high-SES controls, although statistical significance was marginal (LSD, p = .08). That is, the WE group from the low-SES school appeared to lose some of their apparent gains, relative to the C group from the high-SES location.

It can be noted from Figure 3 that the magnitude of change within both WE groups was remarkably similar, with increase in around two points (out of 10), and Hedges' effect sizes in the region of .68 and higher. To further investigate the possible role of prior knowledge we performed median split procedures within each data-set, to contrast students beneath median on pretest with students above the median. This variable did not predict subsequent test gain. Although we cannot assert the null hypothesis, we found that initial scores on T1, which presumably reflect prior knowledge and ability, did not predict the learning gain associated with exposure to worked examples in either experiment. Hence, the notion, that baseline scores did not predict actual gain, appeared valid both across the two studies (comparing E1 with E2) and within each study (per median splits).

3.3.2. Subjective outcomes

As can be seen in Figure 4, the means for self-efficacy within the low-SES school increased virtually to the same level of the C group at the high-SES school (the three means did not differ at T4, with only the WE group from high-SES location being significantly higher). Notably, on indices of cognitive load, the two locations clearly differed, with the high-SES students expressing lower levels throughout. However, by T4, the gap was reduced to the point where the statistical significance between the

WE low-SES group, and the C high-SES group was 12% (i.e. technically not significant, but the magnitude of the gap was closing over sessions).

4. Overall discussion

The goal of the research was to investigate how worked examples could be easily introduced into the classroom at the point (year seven in Australian schools) at which students are have been exposed to proportional reasoning through their curriculum. Our findings suggest that worked examples can have a strong impact on a student's ability to reason proportionally. In one experiment, it was established gains also extended to shifts in students' self-efficacy and self-reported problem difficulty (i.e. reduced cognitive load). Furthermore, the findings show the unitising approach is a valid strategy that can benefit learning. The unitising strategy can be applied to different proportional reasoning contexts. It can provide students with a viable method of reducing complex information to a more manageable form.

The innovatively designed worked examples used in two related experiments exposed students to a series of worked examples. Studying these examples increased students' subsequent test scores a week after this exposure. It was established that such learning gains were evident in samples from schools both high and low in socioeconomic status. In the high-SES school, it was apparent that almost 4 hours of exposure to proportional reasoning problems (i.e. practice alone) did not result in learning gains in students who were not exposed to worked examples. On the other hand, such exposure alone did result in some gains in the lower SES school group, although the magnitude of change was far higher in their peers with access to the worked examples.

These findings are notable in that worked examples were presented to students on paper in the absence of direct teaching. Further, the design of the two projects suggests a viable level of ecological validity. This idea is feasible since, within each location: (a) the context was the classroom, with students anticipating normal lessons, (b) data were collected from students participating in their normal mathematics classes, as described in terms of time and place, with minimal disruption to existing routines, (c) the testing materials and procedures were genuinely "educational" in that the problems used were ones consistent with the curriculum and learning achievements expected of upper primary/lower secondary school age students, and (d) the experimental procedures were carried out by the students' normal teachers.

The first experiment took place in an advantaged school. Hence, it was deemed important to ascertain if such positive effects were achievable in other locations. It was established that learning gains were made by the lower-SES school peers who participated, with those exposed to worked examples making virtually twice the gain associated with the control group (i.e. mean shift of 45 and 21% increases, respectively).

One aspect worthy of note was that the WE group from the low-SES school did appear to respond well to their treatment at the moment of exposure to the worked examples. At the T2 point, these low-SES students actually scored slightly higher than the control group students in the first experiment (respective means of 8.5 and 8.4 on the identical items). In other words, the achievement gap, which was so clearly evident on the baseline measures, was no longer manifest once the low-SES students were given the opportunity to study worked exemplars when placed directly in front of them. This finding suggests that low-SES school students can benefit further from extended exposure to worked examples and explicit strategy training. For instance, fully worked examples could be followed up with partially worked examples, so as to encourage continuity of strategies.

The present experimental design entailed that individual tuition and feedback were not given during the course of the study itself. By design, the two studies were conducted under "test-like" conditions, devoid of interactive teaching opportunities. Further, the follow-up period of seven days may need to be extended to support claims of lasting impacts from the present paper-based treatment alone. Hence, the suggestion is that the impact of exposure to worked examples at one point can be

made more meaningful to the individual student through carefully structured feedback to consolidate gains made from exposure. The provision of initial structured feedback could be of benefit to all students, but especially valuable to novice learners, or as in the present case, to students attending lower SES schools.

Overall, the present findings speak to the issue of teaching proportional reasoning. Despite the fact that this would appear to be a fundamental aspect, a key achievement goal, of the mathematics curriculum, the actual methodologies used to teach proportional reasoning remain remarkably ill-defined. At least within the Australian context, formal ways to teach proportional reasoning skills are not universally accepted. In the classroom, students may be given well-designed problems and encouraged to "*think*", and "*use problem solving skills*". Whilst generic advice often can be helpful, it also can become inherently vacuous, depending upon context and level of student prior knowledge. The findings reported in this paper suggest ways in which teachers can use worked examples to direct students' attention toward readily activated strategies.

4.1. Towards the systematic usage of worked examples in teaching

The findings reported in this paper suggest ways in which teachers can use worked examples to direct students' attention toward critical elements in solving well-structured problems. When teachers are knowledgeable within a curriculum area, they can use their knowledge to develop clear exemplars. Well-designed examples illustrate how complex problems can be recoded and thereby reduced into manageable proportions. We argue that it is important to consider the type and quality of materials when designing worked examples for general classroom use. Carefully designed exemplars will highlight key elements, typically showing a sequence of operations in a step by step fashion. This permits students to more readily grasp how their recently acquired skills can be extended and applied.

We feel it crucial for teachers to collaborate in developing sets of viable worked examples. The underpinning philosophy is that of 'showing how things work'. In a broad sense, such a notion can apply equally well to student learning and teacher professional development. All of us can benefit from carefully identifying the elements that enable one's knowledge to shift from acquisition level to successful applications within a wider range of contexts

Funding
The authors received no direct funding for this research.

Author details
Brendan Bentley[1]
E-mail: brendan.bentley@unisa.edu.au
Gregory C.R. Yates[1]
E-mail: gregory.yates@unisa.edu.au
[1] School of Education, University of South Australia, Adelaide, Australia.

Citation information
Cite this article as: Facilitating proportional reasoning through worked examples: Two classroom-based experiments, Brendan Bentley & Gregory C.R. Yates, *Cogent Education* (2017), 4: 1297213.

References

Artut, P. D., & Pelen, M. S. (2015). 6th grade students' solution strategies on proportional reasoning problems. *Procedia - Social and Behavioral Sciences, 197*, 113–119. https://doi.org/10.1016/j.sbspro.2015.07.066

Ayres, P. (2013). Can the isolated-elements strategy be improved by targeting points of high cognitive load for additional practice? *Learning and Instruction, 23*, 115–124. https://doi.org/10.1016/j.learninstruc.2012.08.002

Capon, N., & Kuhn, D. (1982). Can consumers calculate best buys? *Journal of Consumer Research, 8*, 449–453. https://doi.org/10.1086/jcr.1982.8.issue-4

Cramer, K., Bezuk, N., & Behr, M. (1989). Proportional relationships and unit rates. *Mathematics Teacher, 82*, 537–544.

Dole, S. (2008). Ratio tables to promote proportional reasoning in the primary classroom. *Australian Primary Mathematics Classroom, 13*, 18–22.

English, L. D., & Mulligan, J. T. (2013). *Reconceptualizing early mathematics learning*. Dordecht: Springer Science. https://doi.org/10.1007/978-94-007-6440-8

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analysis relating to achievement*. New York, NY: Routledge.

Hilton, A., Hilton, G., Dole, S., & Goos, M. (2016). Promoting middle school students' proportional reasoning skills through an ongoing professional development programme for teachers. *Educational Studies in Mathematics, 92*, 193–219. https://doi.org/10.1007/s10649-016-9694-7

Jitendra, A. K., Star, J. R., Dupuis, D. N., & Rodriguez, M. C. (2013). Effectiveness of schema-based instruction for improving seventh-grade students' proportional reasoning: A randomized experiment. *Journal of Research on Educational Effectiveness, 6*, 114–136. https://doi.org/10.1080/19345747.2012.725804

Lamon, S. J. (2006). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers* (Vol. 2). Mahwah, NJ: Erlbaum.
Lamon, S. J. (2007). Rational numbers and proportional reasoning: Towards a theoretical framework for research. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629–667). Charlotte, NC: Information Age.
Lesh, R., Post, T., & Behr, M. (1988). Proportional reasoning. In J. Hiebert & M. J. Behr (Eds.), *Number concepts and operations in the middle grades* (pp. 93–118). Reston, VA: National Council of Teachers of Mathematics.
Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction, 16*, 173–199. https://doi.org/10.1207/s1532690xci1602_2
Norton, S. (2005). *The construction of proportional reasoning*. Paper presented at the Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education, Melbourne.
Renkl, A. (2005). The worked-out examples principle in multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 229–246). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511816819
Renkl, A. (2014). *Learning from worked examples: How to prepare students for meaningful problem solving applying science of learning in education: Infusing psychological science into the curriculum*. Washington, DC: American Psychological Association.
Retnowati, E., Ayres, P., & Sweller, J. (2010). Worked example effects in individual and group work settings. *Educational Psychology, 30*, 349–367. https://doi.org/10.1080/01443411003659960
Sowder, J. T. (2007). The mathematical education and development of teachers. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 157–223). Charlotte, NC: Information Age.
Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*, 408–426. https://doi.org/10.1016/j.jecp.2008.11.004
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285. https://doi.org/10.1207/s15516709cog1202_4
Sweller, J., van Merrienboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296. https://doi.org/10.1023/A:1022193728205
van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155–174. https://doi.org/10.1007/s10648-010-9134-7
van Loon-Hillen, N., van Gog, T., & Brand-Gruwel, S. (2012). Effects of worked examples in a primary school mathematics curriculum. *Interactive Learning Environments, 20*, 89–99.

© 2017 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:
Share — copy and redistribute the material in any medium or format
Adapt — remix, transform, and build upon the material for any purpose, even commercially.
The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:
Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.
You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

***Cogent Education* (ISSN: 2331-186X) is published by Cogent OA, part of Taylor & Francis Group.**

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com