Students' Use of Context Knowledge in Interpreting Data

Cynthia Langrall
Illinois State University
<langrall@ilstu edu>

Edward Mooney
Illinois State University
<mooney@ilstu edu>

Nicole Williams
Winona State University
<NWilliams@winona edu>

The written responses of 88 secondary students to a task involving the interpretation of two data sets were examined to study the interplay between students' mathematical/statistical knowledge and context knowledge when interpreting data Findings indicate that most students relied on rudimentary mathematical/statistical procedures although contextual aspects of the task did figure into the work of some students Characteristic patterns of students' reasoning are reported

It is widely acknowledged that proficiency in statistical skills enables people to become productive, participating citizens in an information society and to develop scientific and social inquiry skills Thus, calls for reform in mathematics education have advocated a more pervasive approach to statistics instruction at all levels (e g , Australian Education Council, 1994; Department of Education and Science and the Welsh Office, 1989; National Council of Teachers of Mathematics, 2000) In response to these recommendations, researchers examining the development of students' statistical understanding have begun to produce a coherent body of knowledge that can inform statistics instruction in school mathematics (Ben-Zvi & Garfield, 2004)

Crucial to the development of statistical understanding is the realisation that data are numbers in context (Moore, 1990) By context we mean the real-world phenomena, settings, or conditions from which data are drawn or about which data pertain This definition is consistent with the way the term is used by others (e g , Gal, 2004; Moore, 1990; Pfannkuch & Wild, 2004) with regard to statistics According to Moore, "data engage our knowledge of the context so that we can understand and interpret rather than simply carry out arithmetical operations" (p 96) However, researchers have found that context can create obstacles as well as supports in developing students' statistical understanding (Berg & Phillips, 1994; Mevarech & Kramarsky, 1997)

Because little is known about the interaction between ones' knowledge of context and ability to analyse and interpret data, we have engaged in a series of investigations to examine the ability of context to support or interfere with students' statistical thinking More specifically, we are interested in the following questions: (a) How do students draw on their knowledge of context when analysing and interpreting data? (b) How does a student's level of familiarity with a particular context support or interfere with the ability to analyse and interpret data? The study reported in this paper explored the interplay between students' mathematical/statistical knowledge and context knowledge when solving problems requiring the comparison of two data sets

Theoretical Considerations

The design and analysis of this study was informed by Dapueto and Parenti's (1999) theoretical model for describing the relationship between context and the formation of

mathematical knowledge In adapting aspects of their theory to the domain of statistics, we identified three factors that should serve as considerations when investigating students' statistical reasoning and thinking The first factor involves the students' field of experience, or familiarity with the context of a problem or task The second factor involves the mathematics/statistics inherent in the problem context, that is, whether the context necessitates the use of certain mathematical/statistical knowledge The third factor pertains to the meaningful role the mathematics/statistics plays in understanding or interpreting the problem context From the perspective of the Dapueto and Parenti model, as students build conceptual models or *mathematise* problem situations, they shift (possibly back and forth) between the use of context knowledge and mathematical/statistical knowledge

Other researchers (e g , Pfannkuch & Wild, 2004) have described the notion of shifting or the interplay between data and context According to Pfannkuch and Wild, the ability to integrate statistical and contextual information, knowledge, and conceptions is a fundamental element of statistical thinking They contended that "because information about the real situation is contained in the statistical summaries, a synthesis of statistical and contextual knowledge must operate to draw out what can be learned from the data about the context sphere" (p 20) Similarly, Watson and Callingham (2003) described two essential components for statistical literacy as the "mathematical/statistical understanding of the content and engagement with context in exploiting this understanding" (p 20)

Our work was also informed by Gal's (2004) model of statistical literacy, which describes the types of knowledge (literacy skills, statistical knowledge, mathematical knowledge, context knowledge, and critical questions) and dispositions (beliefs, attitudes, critical stance) that enable a person to "comprehend, interpret, critically evaluate, and react to statistical messages" (p 50) According to Gal, these knowledge bases and dispositions overlap and interactions among them occur as one engages in statistical situations For example, he stated that "if a listener is not familiar with a context in which data were gathered, it becomes more difficult to imagine why a difference between groups can occur, what alternative interpretations may exist for reported findings about an association detected between certain variables, or how a study could go wrong" (p 64) Although Gal's model of statistical literacy was aimed at consumers of statistics, it is also applicable to students in school settings His discussion about interactions among knowledge bases and the notion of taking a critical stance supports our investigation into the interplay between data and context and informed the analysis of data in this study

Method

This exploratory study used a qualitative research design involving a 3-item written assessment followed by individual interviews with selected participants In this paper, we report only on findings from one item of the written assessment, the Bruce Willis Movies Task (see Figure 1)

Participants

The 88 participants in this study were secondary school students who ranged in age from 15 to 18 years They comprised four classes of the same teacher, three pre-calculus and one algebra, and spanned a range of ability levels from average to high Essentially this was a sample of convenience as the assessment tasks required no more than a middle school background in mathematics or statistical reasoning and our interest was in documenting the reasoning of a variety of "typical" students

Task and Procedure

The Bruce Willis Movies Task, as well as the other two items in the assessment, was designed to prompt the comparison of two data sets Data were presented in tables and each set was restricted to five data values to avoid lengthy computations if students choose to examine total values or compute statistical measures Students were permitted to use graphing calculators and were instructed to explain their thinking in writing They were also asked to rate their familiarity with the task context, in this case Bruce Willis and his movies We believed this context would be familiar or of interest to many of the students

Bruce Willis has done a variety of movies including action/suspense movies and comedies Below are the box office earnings for 5 of his recent movies in each genre Is Bruce Willis better off doing action/suspense or comedies? Explain your thinking

Action/Suspense Movies	
Title	**Box Office Earnings**
The Jackal (1997)	$54,911,000
Tears of the Sun (2003)	$43,632,000
The Siege (1998)	$40,932,000
Mercury Rising (1998)	$32,941,000
Last Man Standing (1996)	$18,127,000

Comedies	
Title	**Box Office Earnings**
The Kid (2000)	$69,688,000
Death Becomes Her (1992)	$58,422,000
The Whole Nine Yards (2000)	$57,262,000
North (1994)	$7,138,000
Breakfast of Champions (1999)	$175,370

How familiar are you with Bruce Willis and his movies? Circle one choice:
Not at All Familiar Somewhat Familiar Very Familiar

Figure 1 Bruce Willis movies task

In line with the theoretical perspectives guiding our work, we constructed the task so that students would need to draw on their mathematical/statistical and context knowledge to answer the question of whether Bruce Willis is "better off" doing action/suspense or comedy movies The selection and presentation of data, as well as the ambiguous nature of the question, was intended to prompt students to think critically or take a critical stance (Gal, 2004) with regard to the task

The classroom teacher administered the written assessment as an in-class assignment in each of the four classes Students were told that the purpose of the assessment was to examine how students think about and use data in problem situations They were given a 50-minute class period to respond to the three tasks and most students completed the assessment in that time The Bruce Willis Movies Task was the first item and all students completed this task

Analysis

Students' responses were first organ*is*ed according to their self-reported level of familiarity with the context of the task Then the first and third authors independently assigned one of four codes to each response Three main codes pertained to the evidence of interplay between mathematical/statistical knowledge and context knowledge That is, whether the response reflected consideration of only mathematical/statistical aspects of the task, only contextual aspects, or some interplay between the two Recogn*is*ing that by its

very nature, the data were situated within a context, we considered strictly mathematical/statistical aspects of the task in broad terms that included finding total values or averages or looking for trends in the data Contextual aspects pertained to students' knowledge, beliefs, and attitudes about the context of the task, as well as critical interpretations such as questions about the source or appropriateness of the data, nature of the sample, or need for more information The fourth code was used when students' provided no explanation or responses were uninterpretable All codes were assigned regardless of whether students' interpretations of the data were correct or their reasoning exhibited flaws By way of example, Figure 2 presents sample responses for each of the codes

	Is Bruce Willis better off doing action/suspense movies or comedies?
Mathematics/ Statistics Only	190,543,000 [written under action/suspense data set] 192,685,000 [written under comedy data set] I added up all the earnings for each type of movie, the comedies total was greater than the action/suspense by a couple million Therefore he's better off doing comedies
Context Only	As Bruce Willis gets older and further into his career, action/suspense films will become harder for him to do I think he is better off doing comedies because maintaining the all-time tough guy image for most of one's career in Hollywood will essentially wear out their audience The audience wants to know that he has a soft side too and will be drawn to it
Mathematics/ Statistics & Context	average = 38,108,600 [written under action/suspense data set] average = 38,537,074 [written under comedy data set] According to the average, he is better at comedies, but not by much However, these tables are of his most recent films A more accurate determination can be made by averaging all of his films
Uninterpretable/ No Explanation	He is probably better off doing action/suspense movies

Figure 2 Sample responses for each code

Differences in coding between authors were discussed and full agreement was negotiated (c f , Miles & Huberman, 1994) The final phase of analysis involved examining students' responses within and across categories to discern patterns in the data that characteri*s*ed students' reasoning

Findings

The vast majority of students (93%) indicated they were somewhat or very familiar with the context of this task For most students, however, contextual aspects of the task did not figure into their responses Rather, the majority of students interpreted the data quantitatively, drawing on mathematical/statistical knowledge

		Coded Response				
		Math/Stat Only	Context Only	Math/Stat and Context	No Explanation	
Level of Familiarity	Very Familiar	9	3	4	1	17 (19%)
	Somewhat Familiar	41	4	19	1	65 (74%)
	Not Familiar	5	0	1	0	6 (7%)
		55 (63%)	7 (8%)	24 (27%)	2 (2%)	n = 88

Figure 3 Coded responses by level of familiarity

Characteristics of Students' Reasoning

Mathematics/statistics knowledge only The majority of students' responses were classified as exhibiting mathematical/statistical knowledge only There was no indication that these students attended to any of the contextual aspects of the task although many indicated some level of familiarity with the context Most students described their reasoning in one or more of the following ways: (a) comparison of total box office earnings between genres (appeared in 65% of the responses), (b) notions of consistency of earnings within a genre (29%), (c) comparisons of average earnings between genres (15%), and (d) identification of movie genre that produced the top box office earnings (7%) The following examples illustrate each of these patterns of reasoning:

(a) Comedies, because if you add up the totals for each category [has computed totals for both genres], you can see that comedies made more $ (Student 3Q, very familiar)

(b) Action movies because they make a great amount of money on a continual basis, whereas the comedies had box office incomes that jump around (Student 1BB, somewhat familiar)

(c) Bruce Willis is better off in comedies movies as they tend to make more money on average [has computed averages for both genres] (Student 3C, somewhat familiar)

(d) He is better off doing comedy Not only has he made the most from it [has computed totals for both genres], but his top three earners were comedies Think the guy should take a hint? (Student 1I, not familiar)

Other explanations were reported, but only by one or two students They focused on the box office earnings of the most recent movie, earnings that were considered anomalies (possibly outliers), possible trends in the data within each genre, and percentage comparisons of movies within a genre that performed better than movies in the other genre (e g , three-fifths of the comedies were better than any action movie)

Context knowledge only Few students responded to the task in terms of contextual aspects only Those who did all indicated some level of familiarity with the context Overall, responses reflected the students' beliefs about which genre is more successful in terms of audience appeal or their personal opinions about which type of movie the actor should pursue However, two students drew on their knowledge of Bruce Willis movies not

included in the data sets In some sense, they approached the task with a critical stance albeit supported by personal opinion, as illustrated by the following response:

> He is definitely better off doing action movies because that is what his character is known for Die Hard isn't even on the list which is his most important role in terms of being recogn*i*sed and in terms of financial success His best dramatic role was in Pulp Fiction and that wasn't really a big part So he even knows his place in movies (Student 1P, very familiar)

Mathematics/statistics and context knowledge Aspects of both content and context knowledge were evident in 27% of students' responses, and all but one of these students reported some level of familiarity with the context Two patterns characteri*s*e students' thinking in this category The first pattern, exhibited by half of the students, involved the comparison of total or average earnings for each genre in conjunction with either (a) speculation about the kind of movie preferred by audiences, (b) recognition of other factors that should be considered (e g , information from all movies or production costs), or (c) personal preference for a movie genre Interestingly, although three students acknowledged their own preference for action movies, this did not influence their decision and they each concluded that the data supported the actor's performance in comedy films The following examples show different aspects of this pattern of reasoning:

> I think he is better off doing comedies because over all the comedies has grossed more money than the action/suspense movies Comedies cater to a wider audience more so than action/suspense So I feel he would be better off doing comedies (Student 3P, somewhat familiar)

> Bruce Willis is better off doing action/suspense movies, if you are looking at total gross earnings He made a lot more from his action/suspense movies [no computation recorded] However, there are other factors that could be looked at, including production cost, cost payments, directors, health factors, etc (Student 1X, somewhat familiar)

> Overall, Bruce made more money in his comedy movies [no computation recorded], but I think he does a better job with action/suspense movies just because I see more of a action character than a funny one Bruce would be better off doing comedies because in the chart it shows him making more in that area Which means that people like comedies of his better than action (Student 4J, somewhat familiar)

The second pattern of reasoning, exhibited by one-third of the students, was based on the premise that the box-office earnings for action movies were more consistent or dependable than for comedies This view was coupled with the notion that (a) comedy movies might "bomb" or "flop" or that (b) Bruce Willis is better known as an "action guy" All of these students concluded that the actor was better off with action movies Two examples illustrate this pattern of reasoning:

> Depends on how bad off he is If they choose a comedy for Bruce, he could make more money than an action or he could bomb, but action movies are more dependable to bring in a mid-range earning (Student 2U, somewhat familiar)

> I think that Bruce should do more of the action/suspense movies due to the fact that, for one I like them better, and with the comedies it's kind of hit or miss Sure he has made the most $ out of all of them with the top three comedies, but this genre is nowhere near as consistent as its counterpart (Student 3T, somewhat familiar)

Four students provided multiple reasons that often incorporated aspects of both patterns The nature of these responses is reflected in the two examples below:

> Based on North and Breakfast of Champions, I would have to say that Bruce Willis is successful in comedies only if there is another comedian involved as well However, for the most part, his action/suspense movies are relatively constant in earnings, with the exception of last man standing Based on these observations, I would assume that Bruce Willis is a better Action/suspense actor (Student 2C, not familiar)

Action/suspense While his 3 highest movies are comedies his action/suspense movies always get a good turnout He might make more off of one comedy but he is better off overall sticking with the action ones (Student 2N, somewhat familiar)

Conclusions and Discussion

Although we are limited in the conclusions that can be drawn from this one task, the students' responses do provide some insights into the role that context plays in the interpretation of data With regard to the first question (How do students draw on their knowledge of context when analysing and interpreting data?), the majority of students in this study neglected contextual aspects of the task and relied solely on the use of mathematical/statistical knowledge Of the 31 students who exhibited the use of context knowledge in some way (either alone or in conjunction with mathematical/statistical knowledge), only 11 focused solely on personal beliefs about the situation Most of the others (14 students) interpreted the data within the task context By interpretation we mean that students speculated on audience preference or considered certain movies as "flops" with regard to the data rather than on the basis of only their opinion This conclusion is inconsistent with the findings of others (see Pfannkuch & Wild, 2004) who have reported that students tend to ignore data and instead rely on their beliefs when interpreting data representations These researchers have also noted that students may expect data to support or confirm their personal experiences or knowledge of a situation and this may have been the case with some students in our study The remaining six students who exhibited the use of context knowledge did so by taking a critical stance, usually recogn*i*sing that other factors should be considered in interpreting the data It is disappointing that so few students approached this task from a critical standpoint Although we question whether the structure and presentation of the task inhibited students from critically questioning or evaluating the data, there is evidence in the literature (see Ben-Zvi & Garfield, 2004) that this aspect of statistical thinking may be underdeveloped in statistics education

In terms of the second research question (How does a student's level of familiarity with a particular context support or interfere with the ability to analyse and interpret data?), we were unable to detect any significant relationship between students' level of familiarity with the task and their interpretation of the data The obvious drawback is that this question calls for the examination of students' responses to more than one task However, we are also pursuing an analytical framework that will allow for more structured analysis of the relationship between students' context familiarity and interpretation of data

In general, we believe that the written format of this task limited our ability to fully examine the interplay between students' mathematical/statistical and context knowledge For example, we are unable to discern whether students thought about the context prior to computing averages and then returned to thinking about the context when making sense of the figures they obtained Nor do we know how many students felt compelled to answer the question without questioning the data because this was a class assignment Although the follow-up interviews (not reported here) allowed us to probe students' thinking in this respect, we believe that future research should be classroom based to examine the interactions that occur among students as they engage the data This is the direction we will pursue in our research as we continue to examine how context supports or interferes with students' analysis and interpretation of data

References

Australian Education Council (1994) *Mathematics: A curriculum profile for Australian schools* Carlton, VIC: Curriculum Corporation

Ben-Zvi, D , & Garfield, J (Eds) (2004) *The challenge of developing statistical literacy, reasoning, and thinking* Dordrecht: Kluwer

Berg, C A , & Phillips, D G (1994) An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs *Journal of Research in Science Teaching, 31*, 323-344

Dapueto, C , & Parenti, L (1999) Contributions and obstacles of contexts in the development of mathematical knowledge *Educational Studies in Mathematics, 39*, 1-21

Department of Education and Science and the Welsh Office (1989) *National curriculum: Mathematics for ages 5 to 16* York, UK: Central Office of Information

Gal, I (2004) Adult's statistical literacy: Meaning, components, responsibilities In D Ben-Zvi & J Garfield (Eds), *The challenge of developing statistical literacy, reasoning, and thinking* (pp 47-78) Dordrecht: Kluwer

Mevarech, Z R , & Kramarsky, B (1997) From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions *Educational Studies in Mathematics, 32*, 229-263

Miles, M B , & Huberman, A M (1994) *Qualitative data* Thousand Oaks, CA: Sage

Moore, D (1990) Uncertainty In L Steen (Ed), *On the shoulders of giants: New approaches to numeracy* (pp 95-137) Washington, D C : National Academy Press

National Council of Teachers of Mathematics (2000) *Principles and standards for school mathematics* Reston, VA: Author

Pfannkuch, M , & Wild, C (2004) Towards an understanding of statistical thinking In D Ben-Zvi & J Garfield (Eds), *The challenge of developing statistical literacy, reasoning, and thinking* (pp 17-47) Dordrecht: Kluwer

Watson, J , & Callingham, R (2003) Statistical literacy: A complex hierarchical construct *Statistics Education Research Journal, 2*, 3-46