WHEN STATISTICAL LITERACY REALLY MATTERS: UNDERSTANDING PUBLISHED INFORMATION ABOUT THE HIV/AIDS EPIDEMIC IN SOUTH AFRICA

SALLY HOBDEN
University of KwaZulu-Natal
hobdens1@ukzn.ac.za

ABSTRACT

Information on the HIV/AIDS epidemic in Southern Africa is often interpreted through a veil of secrecy and shame and, I argue, with flawed understanding of basic statistics. This research determined the levels of statistical literacy evident in 316 future Mathematical Literacy teachers' explanations of the median in the context of HIV/AIDS survival times. Drawing on the three-tiered statistical literacy hierarchy proposed by Watson (1998, 2006) and the SOLO taxonomy (Biggs & Collis, 1982), a categorisation framework was constructed. About half the teachers were classified below the level of basic understanding of the median. Misunderstandings included confusion of the median survival time with the maximum survival time, and a failure to consider the spread of the data along with the centre.

Keywords: *Statistics education research; Statistical literacy hierarchy; median; HIV/AIDS data interpretation*

1. BACKGROUND

Statistics, for the general population, is required chiefly for the interpretation of the data they encounter in their daily lives. This data may be in contexts of limited personal importance such as marketing surveys presented in advertisements, or in contexts of great importance such as health matters. In this paper, I begin by describing the in-service teacher re-skilling programme from which the data for this study was drawn, and the HIV/AIDS epidemic in South Africa since these together provide the context for the study. Secondly, the method of data collection is described followed by an explanation of the analytical frameworks used to classify the responses to an item in the module assessment. Thirdly, the results of the analysis are presented. The paper concludes with a discussion of the statistical literacy evident in the teacher responses, and the implications in the context of health literacy for understanding the effects and progress of the HIV/AIDS disease. Two aspects of background are important: the introduction of Mathematical Literacy as a school subject and the consequent need to re-skill teachers, and the HIV/AIDS epidemic in South Africa. These are discussed separately, followed by an explanation of how the two came together in this study.

The high prevalence of HIV/AIDS in Southern Africa has had a devastating impact on communities and families. Data on HIV/AIDS infections and deaths are frequently in the media. Although this information may be in the form of tables of numbers or simple descriptive statistics, it cannot be assumed that people are able to decode and process the information. Information on the HIV/AIDS epidemic is often interpreted through a veil of secrecy, stigma and shame (Duffy, 2006). HIV/AIDS awareness is a focus area of much educational research in South Africa and mathematics education researchers have written about attempts to integrate this into the mathematics curriculum. This has typically involved using the published statistics as a data set for calculating statistical measures and creating graphs or, for example, using the red ribbon symbol to illustrate lines of symmetry (van Laren, 2007). To my mind, although this

Statistics Education Research Journal, 13(2), 72-82, http://iase-web.org/Publications.php?p=SERJ

may raise general awareness it does not enable further understanding of the impact of the epidemic in a powerful way. I suggest that working to interpret the statistical data presented in reports develops the statistical literacy which provides a new lens through which to view the epidemic. In this way, statistical literacy can help develop health literacy. According to Nutbeam, health literacy can be described as "both a goal and an outcome, becoming the currency and capital needed to develop and sustain health" (cited in Ratzan, 2001, p. 210). Peerson and Saunders (2009) argue that health literacy encompasses the capacity to understand and act on information that is central to making health-related decisions, and that it is a higher-order literacy than what they term medical literacy which is more related to understanding a treatment plan, or the required drug dosage. Key to health literacy is the cognitive ability to understand the information, and of particular relevance here, to understand the statistics that are presented in the literature often provided to patients upon diagnosis of a disease.

A feature of the revised curriculum for secondary schooling in South Africa was the introduction of Mathematical Literacy as an alternative subject to the more formal Mathematics traditionally elected by mathematically able learners (Department of Education, 2003). The revised curriculum has a strong agenda of social transformation and redress of the past disadvantage experienced under the apartheid regime. The mathematical literacy subject has the educational and political aims of empowering learners to become responsible citizens, self-managing people and contributing workers. This is very much in line with the definition of quantitative literacy provided by the International Life Skills Survey (ILSS, 2000), which describes a combination of skills, knowledge, beliefs, dispositions, habits of mind, communication capabilities, and problem-solving skills needed for effective engagement with quantitative situations arising in life and work. The key aspect of both mathematical literacy and statistical literacy, as used in this paper, is the ability to use mathematical or statistical knowledge in context and to use this knowledge to make informed judgements.

The boundary between mathematical literacy and statistical literacy is blurred since although only approximately a quarter of the Mathematical Literacy curriculum is explicitly focused on data handling and probability, much of the other content is set in real-life contexts which implicitly require statistical literacy skills such as extraction of data from text, and reading graphs and tables. For the past ten years, statistics has been a formal part of the curriculum from the early grades to Grade 12 Mathematics and Mathematical Literacy. Lajoie (1999) suggests that the introduction of statistics in the K-12 period in the United States had the aim that "students will be better prepared for decision making in the real world" (p. 109).

In response to the need for teachers in the new high school subject of Mathematical Literacy, tertiary institutions offered large-scale teacher development programmes aimed at re-skilling non-mathematics teachers to become qualified Mathematical Literacy teachers. These programmes consisted of two years of part-time study culminating in a professional qualification. The Department of Education funded the programmes and places were offered to selected teachers who had expressed an interest in re-skilling as mathematical Literacy teachers. My work within the re-skilling programme was the development and coordination of a capstone module, "Teaching and Learning Mathematics (with respect to Mathematical Literacy)," taken in the last of the four semesters of the programme. In line with the centrality of current real-life contexts for Mathematical Literacy activities, I decided to use the "Statistics South Africa Statistical release P0302 – Mid-year population estimates 2009" (Statistics South Africa, 2009) for an assessment activity in 2009. The report included data on HIV/AIDS prevalence in the different provinces in South Africa and I thought this was a good opportunity to engage the teachers in making sense of the data. What if working with the numbers could change the way teachers thought about the epidemic? Would they be able to make use of the data to gain insight into the complexities of the HIV/AIDS epidemic issues?

2. RESEARCH DESIGN

This research was part of a larger practitioner research study where I, as the lead teacher of a module, planned to use the data generated in the course of the module activities to understand the learning of the teachers as they re-skilled to become Mathematical Literacy teachers. The two research questions addressed here are: (1) What levels of statistical literacy are evident in teachers' explanations of the median in the context of HIV/AIDS survival times? and (2) what misunderstandings are evident in their explanations?

The participants in this study were practicing teachers of subjects other than mathematics enrolled in a government funded teacher re-skilling programme. Most of the teachers had completed their own schooling and pre-service teacher education in the racially-segregated teacher training colleges of the apartheid era. All the teachers studying the module were potential participants since they completed the assessment activity as a compulsory module requirement. The teachers were given an excerpt from the report to read and study beforehand so that they could obtain clarity on the meaning of terms and the statistical measures used, and for those with English as a second language to have an opportunity to ensure that they understood the text. The assessment was then done individually but with access to any resources teachers elected to bring into the test venue, including textbooks and notes. The teachers had given permission for the work done in the module to be used for my research purposes, and so their answers could be extracted from the assessment scripts after the module was completed.

One item, reproduced below from the assessment activity (Hobden, 2009) provided the data for this study. The teachers were required to explain a sentence containing the word *median*.

> Explain the sentence below ("This release…….for women") in simple English so that a non-mathematical person could understand.
> *Median time from HIV infection to death*
> This release assumed that the median time from HIV infection to death in line with the UNAIDS Reference Group recommendation of 10.5 years for men and 11.5 years for women. (Hobden, 2009, p. 6)

The responses to this question (N = 316) were captured electronically and imported into a qualitative analysis software programme, QSR International's NVivo 10, for coding. Despite the open-ended nature of the question, most of the explanations were single sentences. During review of the data, researchers realized that the coding could be done in two ways– namely, by hierarchical level of statistical understanding, and by categorization of the interpretation of the statement. Consequently, two layers of analysis were completed.

3. ANALYTICAL FRAMEWORK

This section begins with a brief overview of the classification schemes for statistical literacy found in the literature, followed by a presentation of the framework used for this study. Many hierarchical-type structures have been suggested for statistical literacy, or for the understanding of specific statistical concepts (see, for example, Pierce, Chick, Watson, Dalton & Les, 2012; Watson 1998 & 2006; and the literature review of Sharma, Doyle, Shandil & Talaia'atu, 2010). These have many commonalities and it is evident that there is agreement that levels of statistical understanding can be discerned. Watson and Callingham (2003) used archived data from two large-scale research projects to investigate statistical literacy as a unidimensional construct. They suggested six levels of understanding; Idiosyncratic, Informal, Inconsistent, Consistent non-critical, Critical, and Critical mathematical. Investigating professional statistical literacy of teachers, i.e. their ability to interpret the reports provided to the schools containing, for example, assessment data– Pierce and Chick (2011) produced a framework of four nested levels: reading values, comparing values, analysing the data set, and the most encompassing level of interpreting data in local or professional contexts. Groth and Berger (2009) provided another perspective by classifying the responses of pre-service elementary teachers to an item calling for comparison between the mean, median, and mode according to the Structure of the Observed Learning Outcome (SOLO) taxonomy of Biggs and Collis (1982) and the Profound Understanding of Fundamental Mathematics (PUFM) of Ma (1999).

The three–tiered statistical literacy hierarchy proposed by Watson (1998, 2006) provides a simple structure that can be adapted for use in this study. Watson (1998) describes the three tiers of her model as follows:

1. A basic understanding of statistical terminology
2. An understanding of statistical language and concepts as they are embedded in the context of wider social discussion
3. A questioning attitude which can apply more sophisticated concepts to contradict claims made without proper statistical foundation (p. 793)

Only the first two tiers are applicable to this study which is focused on just one item, an item which did not call for the critical judgment characteristic of Tier 3. Within each level, drawing on the work of Biggs and Collis (1982), Watson distinguished three sublevels which allowed for finer analysis: unistructural, multistructural, and relational.

For the purposes of this study, we will use a simplified framework based on the three tiers identified by Watson (1998) and the SOLO taxonomy as adapted by Groth and Bergner (2009). This framework is presented in Table 1. Although some responses are below Tier One, they are classified at the prestructural level within Tier One (1P) which should be understood as a sub-Tier One response – i.e., there is no discernible understanding that the median is a measure of centre or average. Responses that do not meet the requirements of Tier Two as they have no relevant and correct mention of the context (in addition to the statistical understanding coded for in Tier One) could be classified at the prestructural level of Tier Two (2P) but this is superfluous since such response would have been classified at a lower level. Hence while Code 2P is included for completeness, it is not used in practice.

The responses were initially coded using the criteria from an early version of Table 1 as a guide. The criteria were refined as the coding proceeded and clarification was needed. Once all the responses had been coded and the criteria finalised as presented in Table 1, the responses coded within each level were revisited to check the consistency of the code allocations, and to ensure that the highest possible code in the hierarchical framework had been allocated.

4. ANALYSIS

The following two examples illustrate the reasoning used in the coding (see Table 1 for the coding criteria).

Response example 1: *Median refers to the middle; a centre. Those years indicated refer to the middle age of infection. E.g. men infected could be for 21 years and women 23 years from time of infection to death.*

This would be coded 1U. There is *one* relevant aspect, namely the reference to the median as middle or centre. The interpretation of the context is incorrect which precludes coding as 2U.

Response example 2: *This release thinks the middle value (when date is arranged from smallest to largest) of time from HIV infection to death in line with the UNAIDS reference group recommendation of 10.5 years for men and 11.5 years for women.*

This would be coded 1M. There are *two* relevant aspects, the reference to the median as the middle value and the idea of rank order. The lack of interpretation of the context beyond a simple rewriting of the sentence precludes coding as 2M. Full results are presented in Table 2.

An initial scan of the data revealed that certain misunderstandings were repeated. These were tabulated and formed the starting point for a second layer of analysis of the interpretations of the responses. Additional codes were added as they arose. Misunderstandings that occurred five or more times are described in Table 3 where the frequency of each misunderstanding is presented. The contextual implications of each misunderstanding are considered, showing that in this case it really matters in that levels of statistical literacy impact directly on health literacy.

Table 1. Framework for analysis of statistical literacy

Level	Code	Sub-Level	Generic Explanation	Specific explanation
Tier One Understanding statistical terminology (context-free)	1P	Prestructural	Does not address elements of the task. Nothing of relevance is included in the response.	No evidence of centre, middle, average as interpretation of median time.
	1U	Unistructural	Employ single elements of the task. Response shows some understanding and although incomplete, include one aspect of relevance.	Evidence of centre, middle, average as interpretation of median.
	1M	Multistructural	Employ elements in a sequential fashion. Response contains more than one aspect of relevance, but not connected	Mentions middle, centre AND rank order or some aspect specific to median.
	1R	Relational	A unifying explanation of the aspects is provided	Connected and coherent explanation of the median as the middle value of ranked data.
Tier Two Understanding terminology as it occurs in social contexts	2P	Prestructural	Does not address elements of the task. Nothing of relevance is included in the response.	This would be the same as Tier One where context is not correctly alluded to so will not be used.
	2U	Unistructural	Employ single elements of the task. Response shows some understanding and although incomplete, include one aspect of relevance.	Evidence of centre, middle, average number of years as interpretation of median time from infection to death.
	2M	Multistructural	Employ elements in a sequential fashion. Response contain more than one aspect of relevance, but not connected	Mentions centre, middle in rank order of number of years from infection to death i.e. median statistic understood in context
	2R	Relational	A unifying explanation of the aspects is provided	Connected and coherent explanation of the median as the middle value of ranked number of years between infection and death.

Based on Watson (1998) and SOLO Taxonomy (Biggs & Collis, 1982) as used by Groth & Bergner (2009)

5. RESULTS

The teachers showed disappointingly low levels of statistical literacy. Over half of the responses (50.9%) failed to indicate any understanding that the median would be referring in some way to an average or middle value and were classified 1P. In total 42.7% of the cohort were classified at a unistructural level, indicating that they at the very least appreciated that the sentence containing the word *median* had something to do with the centre. More than half of those were able to integrate their understanding of the median as centre with the context leading to their classification at the 2U level. Very few of this cohort of teachers showed understanding of the median as being the centre of ranked data, and only six (less than 2%) were classified at the highest level due to their coherent explanation of the median in the given context. The details of the analysis are presented in Table 2.

In the second layer of analysis, the incorrect interpretations were examined and grouped according to the misunderstanding. In all, nine codes occurring in five or more instances were identified. I have grouped the incorrect understandings under four headings: (a) failure to understand the centre or middle feature implied by the use of median, (b) failure to take the unknown spread of the data into account, (c) going beyond the data, and (d) knowledge of the HIV/AIDS disease interfering with statistical interpretation. The grouped codes together with exemplar responses are presented in Table 3, and discussed more fully below.

(a) The most common misunderstanding or confusion is the apparent understanding of the median time as the actual, or the maximum time rather than understanding that it represents the centre. Such responses which sowed a failure to understand the middle or centre nature of the median were clearly at the prestructural level, since no valid understanding was evident. In the given context, this indicates that the teachers were assuming a much bleaker picture of survival time than is actually being reported. A less-common misunderstanding in this grouping was the assumption that the survival time was at least the reported median, giving a more optimistic view of the survival time. The assumption that the median represented the time when most people would succumb to the disease, the modal time, was not classified as having an awareness of the middle nature of the median as although a recognised measure of centre, it made no sense in this context.

(b) A failure to take the unknown spread of the data into account was evident in 25 responses (8%). To be fair, there was no mention of the spread of the data in the sentence from which the data was drawn, but a conceptual understanding of the median should have led teachers to think about the spread in the given context. The second most common misunderstanding was labelled the half time confusion. Responses in this category showed some inkling of the middle or halfway nature of the median, but regarded the median time as the halfway point between an individual being infected and dying. This appears to take no account of the effect of either the spread or the density of the data. Garfield et al. (2008) remind us that "it is impossible to consider centre without also considering spread, as both ideas are needed to find meaning in analysing data" (p. 188). Linked to this is the implicit assumption of a linear progression of the disease between from HIV infection to death. While it is true that a median survival time of 10.5 years implies that half the HIV infected people would die in less than 10.5 years, it says nothing about the spread of the data. Following a study in Uganda, it was reported that "the median time from seroconversion to AIDS was 9.4 years and from AIDS to death was 9.2 months" (Morgan et al., 2002), which clearly indicates that the disease does not progress in a steady way. Since then the increased use of antiretroviral treatment and better healthcare have improved the prognosis for infected people but the pattern of many years to seroconversion and then less time to death remains. The teachers whose responses were classified as half time confusion showed no appreciation that if the median survival time was 10.5 years, the surviving half of the infected people could possibly die within a year or two, and the spread may not extend another 10.5 years.

Table 2. Levels of statistical literacy of inservice teachers (N= 316)

Level	Exemplar responses coded at this level	*n*	%
1P	This release spells it out that the likelihood of infection is higher for women The number of years one can live from the time of infection. Men can expect to live for 10.5 years and women 11.5 years. If you are a man and get infected at 20 years you could die by age 30.5.	161	50.9
1U	This release estimates the middle number of HIV infection and death in line with the UNAIDS reference group recommendation of 11 years for men and 12 years for women. The middle time from HIV infection to death in line with UNAIDS reference group recommendation of 10.5 years for men and 11.5 years for women	62	19.6
1M	Of all the figures of years collected, they were rearranged in an ascending order and 10.5 years for men and 11.5 years for women were the centre most figures.	8	2.5
1R	Median is the number that is found in the middle e.g. 2 4 6 7 8. 6 is the true median.	5	1.6
2U	Median means middle therefore it means that the middle age for a person to die after infected by HIV/AIDS is 10.5 years for men and 11.5 years for women.	72	22.8
2M	Mostly women live about 11.5 years after the time of infection (HIV) to death. The number at the middle after you have arrange the times that women can live after infected with HIV is 11.5 years	2	0.7
2R	Half the men infected with HIV live for up to 10.5 years from time of infection to time of death. Similarly, half the infected women live for up to 11.5 years.	6	1.9
		316	100.0

(c) In the group of responses which I describe as going beyond the data, incorrect interpretation led to unfounded claims regarding the ages of HIV/AIDS infected people and gender differences in prevalence and death rates. The number of years mentioned in the sentence was misinterpreted in 22 responses as the ages of infected people. We can conjecture that this indicates a language problem and a rather shocking acceptance of the young age at which a person could become infected. The longer median survival time for women was interpreted in 17 responses as meaning that more women than men were infected. In similar vein, eight responses showed an interpretation that the death rate was higher for women. These were unfounded conclusions from this data.

(d) Knowledge of the HIV/AIDS disease interfered with statistical interpretation in 11 cases. The first misunderstanding was that the median time referred to the window period between infection and full-blown AIDS. The second misunderstanding, which I termed the UNAIDS (United Nations Programme on HIV/AIDS) acronym confusion, was evident in those teachers who read the acronym as unAIDS, in other words, either not infected at all, or infected with HIV but not at the stage of full-blown AIDS.

Table 3. Misunderstandings evident in responses

	Node name	Explanation	Example	*n*
a	Maximum instead of median	Explanation implies that the people will die after 10.5 or 11.5 years. The median number of years is taken as the life span.	It means/assumes that if a man is infected he can live with HIV for 10,5 years but women can live with HIV for 11,5 years, therefore men can die first.	79
	Mode instead of median	Indicates that most people die at the given number of years	Median time is the time that is between when you start from the top on the bottom. This means that most people who are infected with HIV die either 10.5 years (men) and 11.5 years (women).	7
	Minimum instead of median	Assumes the given years are AT LEAST how long a person survives	This release assumed that it takes at least 10.5 years for men to die after they have been infected and 11.5 years for women to die from the date of infection	6
b	Half time confusion	Median is an individual's "half time" Projects life to be 23 years Assumes the progression of the disease is linear.	My understanding is that it means from the time a man gets infected until he dies would be (10.5*2) = 21 years because this is the halfway mark	25
c	Age confusion	The median is overlooked and the explanation related to the age (often of infection). Years are taken as the age of the infected people	This shows the number of cases that are infected with HIV are people of the age between 10-12 of both men and women.	22
	Prevalence confusion	Prevalence of infection confused with time to death	It means there are more women which are infected by HIV/AIDS than men.	17
	Death rate confusion	Assumes more men than women die (or vice versa)	It means that to add all province women higher death HIV. Women death all time compared to men. Women death rate is 11.5 years and men is 10.5 years	8
d	HIV to AIDS conversion	Speaks of time as being for AIDS to develop	After being infected as a male you live with virus for about 10 and half years before its fully blown or dying and if you are female you have about 11 and half years before you are infected or dying.	6
	UNAIDS acronym confusion	Acronym understood as not infected with HIV, or as the window period when a person is infected but does not have AIDS	This compares people who are HIV infection and those with UNAIDS.	5

6. DISCUSSION AND EVALUATION

The preceding description of the teachers' interpretations of a sentence related to the HIV/AIDS epidemic raises some points for discussion. These include the limitations of the study, the efficacy of statistics or statistics education courses in developing statistical literacy, and implications for teaching in statistics courses.

It must be noted that the inference of the teachers' levels of statistical literacy was based on just one item, and that they may know more than they wrote down. Classifying the responses was not always straightforward, especially where the teachers were clearly having difficulty expressing themselves in English. Individual interviews would have provided more clarity. This acknowledged limitation to the study is ameliorated by the sample size of 316, a size I consider sufficient to indicate strong patterns of statistical literacy in this cohort of teachers. Analysis of other items in the assessment, beyond the scope of this paper, will add to the data on each teacher and provide deeper insight into their statistical literacy.

A module 'Data Handling in Mathematical Literacy' was offered in the first semester of the teacher re-skilling programme. I was interested to see if they had been examined specifically on the median concept in that module. There was one mention of the median in the examination paper and that was in a question that required the teachers to draw up a marking memorandum for a Grade 11 test. In order to do this, the teachers would have to determine the median mark in each of two (already ordered and with a gap indicating the centre) lists of 30 class marks, calculate the mean and range of the two mark sets and decide which class had done better. This indicates that the teachers who had completed this module had experience with the procedure of determining the median of an ordered list of marks – a task of low cognitive demand. Discussing the performance of the two classes on the basis of two different measures of centre and the range was a higher-level question requiring more insight. Due to administrative problems surrounding the start-up of the teacher re-skilling programme, approximately a quarter of the teachers had not completed the Data Handling module prior to the data collection for this study. While there is no attempt in this paper to link the marks in that module with the level of statistical literacy evidenced by their responses to this single item, it should be noted that 64% of those who had completed the module performed at the prestructural level (1P) whereas 33% of those who had completed the module performed at this level. This is, however, scant encouragement since only a notion of the median as a measure of centre was required to be classified above the prestructural level, and it does raise the question of why so many teachers who had completed the Data Handling module where the median was explicitly discussed produced responses classified at low levels.

The intention of this study was not to highlight the deficits in the teachers whose responses form the data for this study. This would be neither helpful nor fair. The intention is to draw attention to the chasm that exists between the procedural computation of basic statistical measures and the conceptual understanding that enables and informs important opinions. This has been noted by other researchers. For example, Groth and Bergner (2009) conclude after their study on pre-service elementary teachers' conceptual and procedural knowledge of the mean, which revealed rather impoverished understandings, that "a college course in introductory statistics is not sufficient" (p. 58). Despite having completed such a course, students did not appear to have gained deep understanding of the measures of centre. Similarly, following a small-scale study of three elementary school teachers identified as exemplary, Jacobbe (2008) claims that they were unable to make a connection between the procedures for finding the mean and median and the meaning of these measures of centre within specific contexts. "In general, it appears that many students who complete college statistics courses are unable to understand the idea of the mean" (Garfield et al., 2008, p. 190). Cobb and Moore (cited in Groth & Bergner, 2009) note that many introductory statistics courses are driven more by abstract theory than by concrete data.

Why does it seem that typical college statistics courses are unsuccessful in developing the statistical literacy that will be helpful in everyday contexts? Garfield et al. (2008) write that "attaining a deep understanding of seemingly easy statistical concepts is a non-trivial matter, and that there are complex conceptual and procedural ideas that need to be carefully developed" (p. 193). This suggests a style of teaching and learning activities that is not typically found in basic statistic courses or in courses meant to provide students with the requisite content knowledge to teach. Watson and Callingham (2003) suggest that "statistical literacy is not just knowing curriculum-based formulas and definitions but integrating these with an understanding of the increasingly sophisticated and subtle settings within which statistical questions arise" (p. 20). They point out that such high-level skills are unlikely to develop without opportunity to learn through exposure to practice, and this seems to be the key. In South Africa, the school subject Mathematical Literacy explicitly focuses on mathematics in real-life contexts and leaves the theoretical aspects to the core Mathematics school subject. Perhaps, for those for whom statistics is a service course and not an intended line of study, a more explicit focus on statistical literacy is required than is possible in a theoretical statistics course.

7. CONCLUDING REMARKS

This study reveals that a cohort of in-service teachers, mostly with a history of very impoverished schooling, show disappointingly low levels of statistical literacy. Consequently, many were unable to interpret correctly a basic statistical measure, the median, when it is used in the context of the progression of the HIV/AIDS disease. Misinterpreting the median survival time reported has implications in the health literacy domain since one could expect that a person's understanding of the prognosis would impact decisions regarding treatment plans. If the problem is evident in teachers at the end of their two-year teacher development programme (including a module on basic statistics education), one can imagine that it is likely to be a widespread problem in the general population.

Given the fundamental role of teachers in the education of students as future citizens, it is important that both initial and in-service teacher education courses focus as much on statistical reasoning as on the routine calculations. It is only when teachers themselves have high levels of statistical literacy that we can expect them to engender the same in their students.

REFERENCES

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic.

Department of Education (2003). *National Curriculum Statement Grades 10-12 (General) Mathematical Literacy*. Pretoria: Department of Education.

Duffy, L. (2006). Suffering, shame and silence: The stigma of HIV/AIDS. *Journal of the Association of Nurses in AIDS Care*, *16*(1), 13–20.

Garfield, J. B., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). Learning to reason about center. In J. B. Garfield and D. Ben-Zvi (Eds.), *Developing students' statistical reasoning: Connecting research and teaching practice* (pp. 187–200). Dordrecht, The Netherlands: Springer.

Groth, R. E., & Bergner, J. A. (2006). Preservice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning*, *8*(1), 37–63.

Hobden, S. (2009). *Teaching and Learning Mathematics in the FET (w.r.t. Mathematical Literacy); Assignment Three - Real Contexts in Mathematical Literacy*. University of KwaZulu-Natal.

International Life Skills Survey (ILSS). (2000). *International Life Skills Survey*. Ottawa, Canada: Statistics Canada, Policy Research Initiative.

Jacobbe, T. (2008). Elementary school teachers' understanding of the mean and median. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Roundtable Conference*.
[Online: http://iase-web.org/documents/papers/rt2008/T2P13_Jacobbe.pdf]

Lajoie, S. P. (1999). Understanding of Statistics. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 109–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Morgan, D., Mahe, C., Mayanja, B., Okongo, J. M., Lubega, R., & Whitworth, J. A. G. (2002). HIV-1 infection in rural Africa: Is there a difference in median time to AIDS and survival compared with that in industrialised countries? *AIDS*, *16*(4), 597–603.

Peerson, A., & Saunders, M. (2009). Health literacy revisited: What do we mean and why does it matter? *Health Promotion International*, *24*(3), 285–296.

Pierce, R., Chick, H., Watson, J., Dalton, M., & Les, M. (2012, July). *Trialling a professional statistical literacy hierarchy for teachers.* Paper presented at the Mathematics Education: Expanding Horizons: 35th Annual Conference of the Mathematics Education Research Group of Australasia, Singapore.

Ratzan, S. C. (2001). Health literacy: communication for the common good. *Health Promotion International*, *16*(2), 207–214.

Sharma, S., Doyle, P., Shandil, V., & Talakia'atu, S. (2010). Towards understanding models for statistical literacy: A literature review. *Waikato Journal of Education, 15*(3), 115–129.

Statistics South Africa (2009). *Statistical release P0302: Mid-year population estimates 2009*. Pretoria: Author.

van Laren, L. (2007). Using metaphors for integrating HIV and AIDS education in mathematics curriculum in pre-service teacher education: An exploratory classroom study. *International Journal of Inclusive Education*, *11*(4), 461–479.

Watson, J. M. (1998). *Assessment of statistical understanding in a media context.* Paper presented at the Fifth International Conference on Teaching Statistics, Singapore.
[Online: http://iase-web.org/documents/papers/icots5/Topic6w.pdf]

Watson, J. M. (2006). *Statistical literacy at school*. Mahwah, NJ: Lawrence Erlbaum Associates.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3–46.
[Online: http://iase-web.org/documents/SERJ/SERJ2(2)_Watson_Callingham.pdf]

SALLY HOBDEN
University of KwaZulu-Natal
Edgewood Campus
Private Bag X03
Ashwood 3605
South Africa