ASSESSING A STATISTICAL INQUIRY

Noleine Fitzallen[1] Jane Watson[1] Lyn English[2]

[1]University of Tasmania [2]Queensland University of Technology

As statistics education becomes more firmly embedded in the school curriculum and its value across the curriculum is recognised, attention moves from knowing procedures, such as calculating a mean or drawing a graph, to understanding the purpose of a statistical investigation in decision making in many disciplines. As students learn to complete the stages of an investigation, the question of meaningful assessment of the process arises. This paper considers models for carrying out a statistical inquiry and, based on a four-phase model, creates a developmental sequence that can be used for the assessment of outcomes from each of the four phases as well as for the complete inquiry. The developmental sequence is based on the SOLO model, focussing on the "observed" outcomes during the inquiry process.

INTRODUCTION

Contemporary assessment practices acknowledge the advantages of assessing student learning as they work their way through the learning process rather than relying solely on summative assessment conducted upon the completion of a learning sequence. Often termed as "assessment for learning" the evidence gathered is used to monitor student progress and guide the development of subsequent learning activities (Earl, 2003). The goal is to reveal the quality of students' understanding and thinking as well as specific content knowledge and skills development through the integration of assessment into the learning experience. One of the purposes of the shifting emphasis is to support conceptual development of ideas as well as procedural competence. This requires learning sequences to be sustained and ongoing. In statistics education, inquiries that engage students actively in the learning process provide the opportunity for this to be achievable (English & Watson, 2012).

Examples of how to assess the progressive learning outcomes of statistical inquiries are scant. For the most part, assessment reported in statistics education literature is based on statistical literacy rather than actually carrying out a statistical inquiry, involves only part of the practice of statistics, and suggests particular methods such as projects, portfolios, and use of computers (e.g., Bidgood, Hunt, & Jolliffe, 2010; Gal & Garfield, 1997). Research projects have focused on determining the progression of student understanding and application of statistical content or the ability to think and reason statistically, usually accomplished through the evaluation of individual items designed to target particular statistical concepts. This type of research has led to the development of hierarchies of learning that characterise student understanding at different levels, such as a statistical literacy framework (Watson, 2006). There are, however, few examples of research that involves the assessment of student understanding as they work systematically through a statistical

2015. In Beswick, K., Muir, T., & Wells, J. (Eds.). *Proceedings of 39th Psychology of Mathematics Education conference*, Vol. 2, pp. 305-312. Hobart, Australia: PME.

inquiry.

TOWARDS A STATISTICAL INQUIRY FRAMEWORK

Fundamental to an inquiry-based pedagogy is the need for teachers not only to allow students to construct their own learning but also to support and scaffold that learning (Bell, Urhahne, Schanze, & Ploetzner, 2010; Makar, 2012). This is fostered through the application of inquiry frameworks that guide the implementation of a series of learning activities. Common to the many inquiry frameworks described in the literature are orientation and questioning processes in the beginning, followed by processes of investigation, and finalised with activities that demand students draw conclusions and evaluate findings (Bell et al.). These processes are exemplified in a traditional science inquiry that involves students working through a sequence of question, predict, experiment, model, and apply (White & Frederiksen, 1998).

Statistics education research offers some examples of frameworks for describing statistical thinking and reasoning that encompass the notions of inquiry. An extensive four-dimensional model proposed by Wild and Pfannkuch (1999) and elaborated on by Pfannkuch and Wild in 2004, is based on the way statisticians work and think statistically and can be applied to the way in which students engage in statistical investigations. It includes: Dimension 1: The investigative cycle, Dimension 2: Types of thinking; Dimension 3: The interrogative cycle; and Dimension 4: Dispositions.

Dimension 1 is related to the thinking processes employed when working through a statistical investigation. This involves posing a question, planning an investigation, collecting data, analysing data, and drawing conclusions. Dimension 2 is related to general and particularly statistical thinking. Wild and Pfannkuch (1999) posit that the types of thinking in this dimension are "the foundations on which statistical thinking rests" (p. 227). Dimension 3 adopts a cyclical process of data interrogation that involves thinking critically about the data in order to distil and encapsulate ideas and information. Dimension 4 includes the personal qualities, dispositions, and habits of mind employed when working with data. These dimensions underpin the way in which people should work statistically but are not all transferred easily to teaching and learning contexts.

More recently, Allmond, Wells, and Makar (2010) have provided guidance on how to develop learning experiences with a mathematical inquiry focus. Their framework encompasses the need to make connections to the context of a problem and to recognise the purpose of investigating a problem; to plan an investigation that provides the evidence needed to answer the problem; to draw on a range of mathematical concepts and skills to collect, represent, and interpret data; and to communicate and justify findings to an audience. Their model includes the following phases:

- Discover: Situating a question within a context and understanding its purpose.
- Devise: Planning an investigation.

- Develop: Engaging in mathematical reasoning.
- Defend: Communicating and justifying a conclusion.

The mathematical inquiry framework developed by Allmond and her colleagues is generic and can be applied to any mathematical investigation. It does not, however, provide specific guidance for designing or implementing a statistical inquiry.

Similar to other inquiry frameworks, the *Model of Statistical Investigation* (Figure 1) developed by Watson (2009) starts a statistical investigation with a question set in a context. The question sets the scene for an inquiry and draws in the context of the data. The Data Collection step provides the data that can be represented in a number of forms – numerical, pictorial or graphical. The data are often reduced using statistical calculations of measures of centre or graphical representations such as a box-and-whisker plot. These representations or measures are then used to make inferences about the data that answer the question posed in the initial stage of the inquiry. Part of the inference step is recognising the level of uncertainty associated with the conclusions drawn. The steps encompass the aspects of working statistically detailed in the *GAISE Report* curriculum framework (Franklin et al., 2007), particularly, "actively collecting, organising, summarising and interpreting data" (p. 63). In addition, Watson's model recognises the importance of context, the notion that not all conclusions can be made with the same level of confidence, and the underpinning idea that variation is fundamental to all statistical inquiries. This model provides a comprehensive view of what students should do as part of an inquiry.

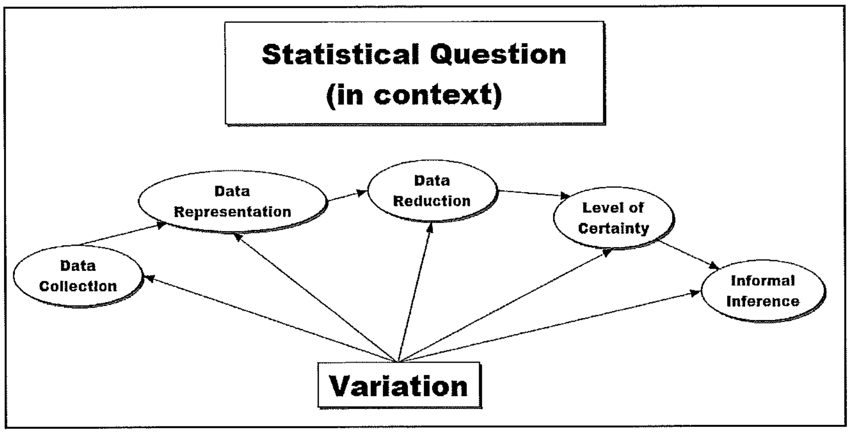

Figure 1. Model of Statistical Investigation (Adapted from Watson, 2009, p. 91).

Acknowledging the potential complexity of the model in Figure 1 for the classroom and in accordance with *GAISE*, an inquiry is summarised in four phases: (1) Pose questions, (2) Collect data, (3) Analyse data, (4) Make decisions.

These four phases constitute a statistical inquiry framework that is more applicable in statistics education than generic inquiry frameworks or those developed for science investigations. At the classroom level it has the potential to support teachers to plan a meaningful sequence of learning that can be communicated easily to students, who can keep track of their progress throughout the inquiry by relating activities undertaken at any stage of the inquiry to the phases of the framework. From a research perspective, it details specifically phases of a statistical inquiry that can be evaluated and interrogated individually to ascertain student understanding at particular phases of an inquiry. It encapsulates the notions of statistical thinking and reasoning highlighted by Pfannkuch and Wild (2004) and the process of inquiry advocated by Allmond et al. (2010).

IMPLEMENTING A STATISTICAL INQUIRY

It is essential for strategies developed for the assessment of a statistical inquiry to accommodate the way in which an inquiry is implemented. This process is not necessarily linear and entry into an inquiry may occur at any of the four phases depending on the background provided to students. This flexibility offers the opportunity to scaffold student learning and stage the development of the necessary skills and strategies for each of the phases before they are required to work through a complete inquiry, which includes working through the four phases sequentially. The advantage of using a staged approach is that skills developed during one phase may be consolidated with the skills developed from another phase (Watson & Fitzallen, 2010), thereby building students' capacity to initiate and work through the full inquiry process.

The implication for using a statistical inquiry either completely or partially is that teaching and learning and assessment practises need to accommodate both scenarios. When teachers and researchers use activities that focus on one of the statistical inquiry phases, they need to be aware of the elements of that phase so that they can provide the support needed for students to bring the ideas from each of the elements together to develop an understanding of the learning outcomes associated with that particular phase.

The way in which elements of understanding can underpin a concept is exemplified in a general developmental model of graph creation (Watson & Fitzallen, 2010), which includes three hierarchical sequences of learning development: the concept of graph, the ability to create or choose appropriate graphs, and informal decision making from graphs. In the context of graph creation and interpretation the second sequences constitutes two parallel sub-sequences: one for when more than one attribute is involved and another for when a large data set is used.

DESCRIBING LEVELS OF UNDERSTANDING

An essential component of the assessment process is having a structured sequence of the expected learning and its outcome. A useful model is the Structure of Observed Learning Outcomes (SOLO) of Biggs and Collis (1982). In assessing learning outcomes, such as related to a statistical inquiry, the focus is on what is observed during the process rather than responding to a test item at a later time.

A neo-Piagetian model, SOLO includes multiple modes of functioning, of which the concrete symbolic is of interest here because of the symbolic learning that takes place in schools based in empirical elements and concrete materials. Within the concrete symbolic mode, learning sequences can be identified in a hierarchy described as prestructural (P), unistructural (U), multistructural (M), and relational (R). In terms of the elements provided as part of the learning tasks, there may be none employed (P level); single elements may be used but are totally unrelated to each other (U-level); several separate elements may be employed in a sequence (M-level); or all of the elements may be combined in an integrated fashion showing their relationship to produce a conclusion (R-level). Once the result of a particular learning sequence is consolidated, it may in turn provide an element for a higher order sequence for which it is an essential ingredient. Watson and Fitzallen (2010) illustrated this for the concept of average and the development of graph understanding and its application to graph interpretation and decision making. Other characteristics of the U-M-R levels include the potential lack of recognition of conflict or identification of contradiction at the U-level, their recognition but lack of resolution at the M-level, and their resolution at the R-level should they arise. Conflict or contradictions arise when decisions made and ideas expressed by students are incorrect or there is a mis-match of ideas and information.

As a starting point for assessing learning outcomes from a complete statistical inquiry it is suggested that there are U-M-R sequences associated with each of the four phases of an inquiry. Each of these phases, when complete, provides an element for a U-M-R sequence that describes the development of understanding the practice of statistics. A general developmental sequence is seen in Figure 2. It applies either to one of the phases or to the complete inquiry.

	Successful Learning Outcome			
Relational level	Combines all elements in an integrated fashion to achieve the outcome; resolves any conflicts/contradictions recognised.			
Multistructural level	Links several elements in sequence; may recognise but not resolve conflict/contradictions.			
Unistructural level	Use single elements unlinked; does not recognise conflicts/contradictions.			
	Element 1	Element 2	Element 3	Element4*

*There may be more than 4 necessary elements.

Figure 2. General developmental sequence for a phase of a statistical inquiry.

The consolidation of each phase of the inquiry becomes a new Element to be employed in a subsequent phase or complete inquiry. Table 1 suggests the elements that are likely to be employed in the phases of an inquiry. Although acknowledging student achievement could occur at any of the levels of the developmental hierarchy, the relational level is the desired level of achievement for the targeted learning outcome.

Inquiry Phase	Elements

Pose questions	Context, Population, Type of measurement, Attributes
Collect data	Question, Type of data, Instruments, Sample Size, Variation
Analyse data	Question, Data, Graphical representation tools, Data reduction tools
Make decisions	Context, Question, Analysis, Uncertainty, Interpretation

Table 1. Elements required for the 4 phases of a statistical inquiry.

To exemplify the relationships in Figure 2 the inquiry phase of Pose questions is illustrated in Figure 3 with the Elements suggested in Table 1. The examples are from Year 6 students involved in posing and refining questions within the context of a claim that "athletes are improving over time" (English & Watson, 2012). Not all students chose the same sport or sporting event, measurements, or time frames. This resulted in many different examples at each of the developmental levels. Also varying among the students' responses was the way in which they incorporated the different elements identified.

	Pose Questions (within a context) e.g., "A claim that Athletes are improving over time"			
R-level	All elements integrated; no conflict or contradiction; e.g., "Are the times of the gold medal 100m men sprinters in seconds generally improving over the years of each Olympic Games?"			
M-level	Several elements in sequence; recognise but not completely resolve of conflict or contradiction; e.g., "Are people who sprint 100m at Olympics improving over their career?"			
U-level	Single elements, unlinked, unrecognised conflict or contradiction; e.g., "In what age group do 100m men's athletes win gold?"			
Elements	Pick sporting event; e.g., Olympic games (Context)	Specify athletes; e.g., men's100m sprint (Population)	Specify dates; e.g., 1896-2012 (Type of measurement)	Identify measurement; e.g., winner's time (Measurement criterion)

Figure 3. Examples of student responses across the developmental sequence for the *Pose questions* phase of an inquiry about athletics performance.

Next is a culminating sequence, which uses the outcomes of the four phases as elements for a complete statistical inquiry (Figure 4). This sequence recognises the necessity to integrate all phases of the inquiry, depending on the task set, and may represent thinking moving from the concrete symbolic to the formal mode of the SOLO model (Biggs & Collis, 1982).

	Complete statistical inquiry			
R-level	Integrates all elements; e.g., includes all 4 Elements accurately combined with uncertainty recognised in the Decision			
M-level	Links several elements in sequence; e.g., sets up Question with Analysis and Decision without Data or recognition of Uncertainty			
U-level	Single parts of the inquiry unlinked; e.g., only discusses Analysis			
Elements	Pose questions	Collect data	Analyse data	Make decisions

Figure 4. General developmental sequence for a complete statistical inquiry.

CONCLUSION

Assessment of the learning outcomes generated by students completing multi-staged statistical inquiries is complex. It needs to encompass evaluation of both the understanding of statistical content and the application of statistical processes. Presented in the previous sections is an example of how assessment can be integrated into the learning process when students conduct statistical inquiries. The identification of the phases of an inquiry and the characterisation of the elements that make up each of the phases, together with the application of SOLO (Biggs & Collis, 1982) to describe the progression of learning within each phase and across a complete inquiry, provide assessment constructs that have the potential to be used to not only determine students' level of achievement at various stages throughout a statistical inquiry but also support curriculum planning (Earl, 2003). As a teaching point, the developmental nature of the sequences sets up the importance of students reaching the R-level as a desired end point for each individual phase and complete inquiry. Supporting teachers to be cognisant of whether students have attained understanding at that level may also result in the utilisation of activities that provide the opportunity to achieve at that level. Specifying the statistical knowledge and skills associated with the elements of each phase signposts for teachers aspects of the statistics curriculum that need to be emphasised and highlighted. The way in which the outcomes from one phase are consolidated in subsequent phases has the potential to facilitate students' transitions across phases (Watson & Fitzallen, 2010). This also has the potential to support students to develop the skills and knowledge necessary to have the capacity to progress independently through a complete statistical inquiry, a focus of future research, which will involve investigating ways of translating the assessment constructs described in this paper into classroom assessment tools (English, Watson, & Fitzallen, 2015).

Acknowledgement

This research is supported by a Discovery Project Grant (DP120100158) from the Australian Research Council.

References

Allmond, S., Wells, J., & Makar, K. (2010). *Thinking through mathematics: Engaging students with inquiry-based learning (Books 1-3)*. Melbourne, VIC: Curriculum Press.

Bell, T., Urhahne, D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: Models, tools, and challenges. *International Journal of Science Education, 32*(3), 349–377.

Bidgood, P., Hunt, N., & Jolliffe, F. (2010). *Assessment methods in statistical education: An international perspective*. Chichester, UK: Wiley.

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

Brady, L., & Kennedy, K. (2012). *Assessment and reporting: Celebrating student achievement*. Frenchs Forest, NSW: Pearson.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning. Experts in Assessment series*. Thousand Oaks, CA: Corwin Press Inc.

English, L. D., & Watson, J. M. (2012). *Statistical literacy in the primary school: Beginning inference*. [DP120100158] Retrieved from http://www.arc.gov.au/pdf/DP12/DP12_Listing_by_all_State_Organisation.pdf

English, L. D., Watson, J. M., & Fitzallen, N. (2015). *Modelling with data: Advancing STEM in the primary curriculum*. [DP150100120] Retrieved from http://www.arc.gov.au/media/feature_articles/Dec2014_Lyn_English.html

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework*. Retrieved from http://www.amstat.org/education/gaise/

Gal, I., & Garfield, J. (1997). *The assessment challenge in statistics education*. Amsterdam: IOS Press & The International Statistical Institute.

Makar, K. (2012). The pedagogy of mathematics inquiry. In R. M. Gillies (Ed.), *Pedagogy: New developments in the learning sciences* (pp. 371-397). New York: Nova Science Publishers.

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. 17-46). Dordrecht, Netherlands: Kluwer Academic Publishers.

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

Watson, J. M. (2009). The development of statistical understanding at the elementary school level. *Mediterranean Journal of Mathematics Education, 8*(1), 89-109.

Watson, J. M., & Fitzallen, N. E. (2010). *Development of graph understanding in the mathematics curriculum. Report for the NSW Department of Education and Training*. Sydney: *NSW Department of Education and Training*.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-265.