STUDENTS' EMERGENT ROLES IN DEVELOPING THEIR REASONING ABOUT UNCERTAINTY AND MODELING

Dani Ben-Zvi and Keren Aridor
The University of Haifa, Israel
kerenaridor@gmail.com, dbenzvi@univ.haifa.ac.il

Roles that students take in solving problems can help in guiding and scaffolding their learning and meaning making. We present a case study – part of a UK-Israel research project – that focuses on the emerging roles spontaneously developed by Israeli eighth-grade students (14 years old) in solving a scientific-statistical inquiry task using TinkerPlots2. The task integrated four design approaches: Exploratory Data Analysis, Active Graphing, modeling, and gaming. We examine how this task design played a role in this emergence of students' roles and how they respectively adopted perspectives on uncertainty and modeling. Implications of the findings are discussed.

INTRODUCTION

The recent decades have seen a change in the teaching and learning of statistics in content, pedagogy, and technology (Moore, 1997). The current educational focus is on developing students' statistics literacy, reasoning and thinking rather than transmitting statistical procedures, computations and graphing skills (Garfield & Ben-Zvi, 2008). Statistics education researchers therefore pay growing attention to students' emergent statistical reasoning and to the conditions that support it. This educational reform is supported by the application of new approaches to teaching statistics, such as Exploratory Data Analysis (EDA, Tukey, 1977) and Active Graphing (AG, Ainley, Nardi & Pratt, 2000). One key aspect that received less attention in statistics education is the *roles* that students assume in solving statistical tasks. These roles can become learning scaffolds by providing impetus and structure to students' actions and considerations, collaboration and building new knowledge and skills (Jahnke, 2010).

This study is part of a UK-Israel research project[1] (Ainley, Aridor, Ben-Zvi, Manor & Pratt, 2013). In this paper we demonstrate how an integrated task design encouraged the students to take different roles, and how these roles played a role in their understanding the task and developing statistical reasoning. We discuss roles in learning and present the design of the *Angry Emu* task that combine EDA, AG, modeling and gaming approaches. We then focus on several episodes that illustrate the different roles that two eighth-grade students adopted, and how these roles affected their reasoning about uncertainty and modeling. Finally, we discuss implications of this case study.

SCIENTIFIC BACKGROUND

Roles in learning. In social contexts, roles are prescribed functions and responsibilities providing guidance for individual behavior and regulations for group interaction and collaboration. They can be prescribed implicitly or explicitly (Hare, 1994) and may refer to sets of interaction patterns with different functions, meanings or purposes (Jahnke, 2010), or a set of activities performed by individuals (Goffman, 1959). The conceptualization of roles in educational contexts is diverse (Mudrack & Farrell, 1995). Some researchers view student's roles as scaffolds for learning processes that aim at building new knowledge and acquiring cognitive and collaborative skills. Three levels of roles were categorized by Strijbos and De Laat (2010): the micro, meso and macro levels. Each of these levels has a different impact on the collaborative process of the learning group and of the individual learning process in it. At the micro-level (role as a task), a role consists of a task specifying a single collaborative product-oriented or process-oriented activity. At the meso-level (a role as a pattern), the role consists of a pattern of multiple tasks, focused on the product, process or a combination. At the macro-level (a role as stance) the role consists of a stance comprised of an individuals' participative style based on their attitude toward the task and collaborative learning. This framework suggests another related distinction between scripted and emerging roles. Scripted roles are prescribed by the teacher or the task designer, and are usually roles at the micro and meso levels. Emerging roles are spontaneously developed by students and are usually roles at macro levels.

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

Roles can encourage interdependence and individual accountability (Slavin, 1995), increase group members' commitment to the tasks and goals of the group and awareness to its performance, and subsequently increase group functioning (Strijbos & De Laat, 2010). Both emerging and scripted roles provide structure to the activity, which can give the students a sense of security needed to concentrate on the task and enable the collaboration needed to complete it (Morris et al., 2010). In collaborative learning, roles are most relevant for distributing, coordinating and integrating sub-tasks in order to attain a shared goal (Strijbos & De Laat, 2010).

In the current study we refer to aspects of students' *emerging roles at the macro level*, which shaped their participative style, attitude and understanding the task, and their actions and reasoning. We demonstrate the way these roles supported or hindered the development of statistical reasoning, and how the task design has played a role in the students' role selection. The task design included four approaches: EDA, AG, modeling and gaming. Each of them is explained briefly.

Exploratory data analysis (EDA), developed first by Tukey (1977), is an accepted way of approaching the analysis of data (Biehler, 1990). EDA can help develop students' statistical reasoning through collecting, organizing, describing and analyzing of data, with emphasis on simple sense making tools and visual representations, usually with the aid of technology, for interpretation, analysis and inference (Moore, 1997). Making sense of data in its context includes clarification of the reasons and purposes for collecting the data, recognizing patterns and trends, formulating conjectures, explaining and drawing informal inferences (Pfannkuch, Rubick, & Yoon, 2002). These complex processes require the ability to interpret graphs, which is challenging for students (Ainley, Pratt & Nardi, 2001).

Active Graphing (AG) is a pedagogic strategy developed by Ainley and Pratt to support students' understanding of graphs. Students are involved in an experimental process requiring them to use graphs (created utilizing technology) to solve meaningful problems. During this process students need to be aware of the dependence relationship between variables, collect, organize and analyse small amounts of data in order to decide about their next step in the experiment. Ideally, this process includes several iterations of data collection and graphing. The necessity of making sense of graphs during this process encourages student to capture the graph as an analytical tool, rather than just as a tool for data representation (Ainley et al., 2000). In the current task design, we used modeling as a bridge to integrate EDA and AG (Ainley et al., 2013).

Modeling is a key component and tool in science and probability. Integration of probability simulation with data analysis tools, for instance in TinkerPlots2 (TP2), can allow students to explore the connections between *data and chance* while investigating real phenomena (Konold, Harradine & Kazak, 2007). The use of computer modeling and probabilistic simulations of many samples can introduce students to central ideas, such as thinking about data as comprised of signal and noise, sampling as an iterative process used to evaluate conjectures by collecting and analyzing data to improve model and (informal) inferences, reasoning about distributions and the law of large numbers. Finally, the fourth design component was gaming.

Gaming. Game design involves a variety of knowledge and skills, not only in exploring the space of a set of rules, but mainly learning to understand and evaluate a game's meaning as the product of relationships between elements in a complex system (Salen, 2007). Therefore, game design can encourage meta-level reflection on the skills and processes that the designer undergo during its construction: a reflection-in-action process including an iterative sequence of modifications of the rules and the behavior of the game components (Schön, 1983).

METHOD

In this paper we focus on the work of two Israeli students. They handled the *Angry Emu* task – an authentic scientific inquiry with clear purposes, which involved them in planning, prediction, explanation, statistical modeling and informal statistical inference. The *research questions* are: a) How did the task design – based on EDA, AG, modeling and gaming approaches – play a part in students' spontaneously taking roles? and b) how did these roles play a part in promoting or hindering students' statistical reasoning?

The Angry Emu Task. The children were asked to help Rovio (developer of the popular Angry Birds game) develop a computer game, The Angry Emu, in a way that will resemble the real movement of a toy bird, that cannot fly in the air and moves only horizontally. The students were asked to prepare a data-based recommendation, as a result of studying different situations, making conjectures and testing them with the toy bird. To do this, they needed to follow several stages: discuss expectations about the motion of the Angry Emu; collect data about the bird's motion, record and graph it in TP2; discuss the graphs in light of the expectations and predict the results of further experiments; build a Model in TP2 that represents the bird's movement; sample from the model, draw sample graphs and compare them to the real data plot; evaluate the model and improve it; make inferences about the bird's motion and express their certainty level in these inferences; and prepare a report to Rovio. This process was designed to encourage students to express and evaluate uncertainty and to grow key statistical ideas. We added a gaming component to make the task more challenging for the boys by providing rules, such as, "if the Angry Emu arrives at the target range – the player gets the highest score, 100 points."

Participants. The participants are Shay and Dor, a pair of academically successful and articulate 14 year-olds (grade 8) from a science focused private school in Israel. Shay and Dor participated in three cycles of the Connections project (Gil & Ben-Zvi, 2011). They experienced drawing informal inferences from samples of increasing sizes using TP1 in grade 5, engaged in real world data and model-based investigations using TP2 in grade 6, and drew informal inferences from samples in an "integrated approach" using TP2 in grade 7 (Manor, Ben-Zvi & Aridor, 2013).

The Actual Learning trajectory. Shay and Dor's participated in three 90-minute episodes and an interview (for more details see Ainley et al., 2013). In Episode 1 they were introduced to the task, planned a launching stand and experimented a little with it. In Episode 2 they planned and constructed the launching stand with the Emu's toy. They considered the connection between the real experiment and a computer game, the rubber tension, the distance of the launching stand from the wall, the position of the target point (hundred points), the number of stretching levels and repetitions required to infer about a range of values. They decided to set five levels of stretch, but collected only six measurements for stretch three – the "average stretch." Based on these data, they fixed the distance of the launching stand from the wall and the position of the target point, created an initial model in TP2 and predicted distances for stretches one and three. In Episode 3 the boys completed their TP2 model for distances of stretches one, three and five, generated a random sample (n=10) from the model, compared it to the collected data, and estimated their level of confidence in the model and the chosen sample size. They then collected four additional measurements for stretches one, three and five and returned to evaluate and improve their model. At the end, the students described how the results would be introduced to the manufacturer.

Data collection and analysis. Episodes 1 to 3 were attended by two researchers who frequently asked the students to clarify their reasoning. All episodes were entirely videotaped and computer screens were captured using Camtasia. The videos were carefully observed, transcribed, translated from Hebrew to English, and annotated for further analysis of the the roles students took and the ways these roles took part in their statistical reasoning. We used an interpretive microgenetic method (e.g., Siegler, 2006) taking into account verbal, gestural, and symbolic actions within the situations in which they occurred. Interpretations were discussed by the UK-Israel research team, until consensus was reached.

RESULTS

We identified two roles in these data: the *gamer* (Dor) and the *researcher* (Shay). The *gamer* aimed to design an interesting and challenging game experience, regardless of the actual movement of the bird. The *researcher* aimed to better understand and describe reality in order to design as realistic a game as possible. The following results illustrate: a) the different roles the boys took influenced by the four task design approaches, and b) the ways these roles played a part in how they approached experiment and modeling and expressed their reasoning about uncertainty.

Plan, Experiment and Model (Episode 2)

While reading the task, Shay immediately pointed out[2]: "In reality, there is a chaotic system, which means that even if you launch the bird in the exact same angle, in the exact same power, and even if a robot does it, you still won't get the same result" [line 35][3]. In taking the role of the researcher, Shay was genuinely interested in studying the toy's real movement in order to develop a game that would replicate it adequately. He pre-planned his report to Rovio: "I can tell

Rovio a range of numbers… the mean, the bigger chance. What they can do is to write a program that in fact tosses random numbers and also results in numbers with a bigger chance to be drawn" [49]. Shay's chosen role thus encouraged him to use statistical concepts he knew from previous years, such as distribution, center, variability, randomness, certainty and chance. He expected to control the uncertainty by creating a linear model of the motion coupled with local chance frequencies. Dor's intentions were however different:

64 D *In my opinion, if the computer screen is the scale of the game, then the rubber band should be stretched approximately up to here [pointing to a position on the floor]*[4] *…*

78 D *What we need to do is divide it into lines [parts] to which the user can stretch the rubber band up to… and to do several experiments on every line, and decide how many experiments.*

79 D *And then, we actually enter the tension of the rubber band into the [computer] program in relation to where it [the bird toy] reached.*

Like Shay, Dor's previous experiences with EDA investigations encouraged him to first take the role of the *researcher*. He planned an experiment to make data-based inferences and expressed his understanding of uncertainty by the need for repetitions [78]. However, at this point his emergent role as a *gamer* quickly took over, and his main goal was to make an engaging game, even if this means changing or ignoring reality. Influenced by the game illustration, he planned to build the launching point based on a hunch rather than a systematic experiment [64]. Although he acknowledged the existence of uncertainty, he chose to ignore it in pre-planning the report to Rovio and modeling the bird's movement in a deterministic manner [79].

After Shay and Dor collected data of six launches of stretch three (Figure 1), they modeled the predicted bird movement in TP2. Shay first drew a probability distribution of distance ranges for stretch three (Figure 2) based on the data they collected. Shay added: "I don't want to tell the computer to draw between this number and that number. I [rather] tell it: There is 80% that it [the distance] will come out between this number and that number, 15% that it will come out between this number and that number, or between this number and that number, because they are on both sides [of the mean], and 5% that it will come out such and such, okay? And then it [the computer] draws randomly in such a way that 80% will come out the 80%, 15% will come out the 15% and 5% will come out the 5%" [340]. The researcher's point of view and the need to model the phenomena encouraged Shay to construct a data-based probability model quantifying uncertainty.

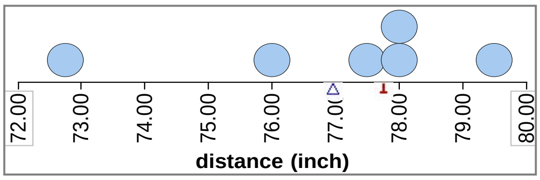

Fig. 1: Data collection of stretch 3.

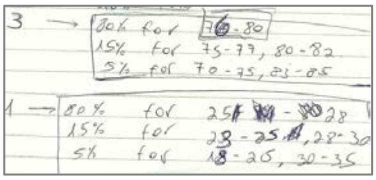
Fig. 2: Shay's probability distributions.

Dor the "gamer" focused on designing an exciting game and simplifying the report by trying to use the mean distance for each stretch. However, the requirement to build a model made him suggest a distribution of possible distances for each stretch (e.g., stretch one's distribution is presented in Figure 3). It seems that the modeling approach integrated into the task design, added complexity to the different roles the boys took and to their reasoning about uncertainty. Moreover, they were not always consistent with the roles they took. For example, when they continued constructing their model, they argued about the way stretch five should be modeled.

81 S *I would agree that this is how we should program in the game. I wouldn't agree with you that this is what happens in reality.*

82 D *But we want to find out the situation in reality.*

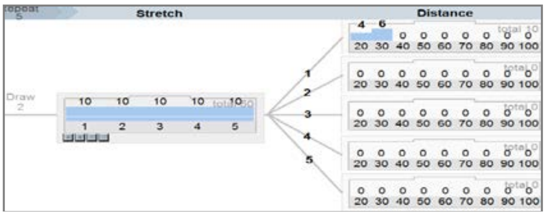

Fig. 3: Dor's model of stretch one.

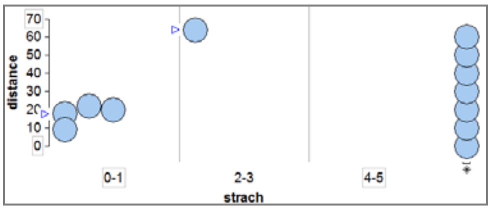

Fig. 4: Second data collection.

Shay explicitly expressed the differences between the two roles [81], while Dor switched in this debate to the researcher's role [82], although he mostly focused during the activity on creating an interesting game rather than a realistic one.

Collect Additional Data and Remodel (Episode 3)

After the boys created their first model, they took random samples from the model to decide about the necessary sample size and the number of repetitions. They then collected more data in order to verify and improve their model. They used the AG approach to measure four distances of each stretch. As soon as they entered the first measurement of stretch three into TP2 (Figure 4), Shay noticed an interesting trend.

208 S *Just a moment. I suddenly have a hypothesis. I think that the difference between [stretch] one and [stretch] three is much larger than it is between [stretch] three and [stretch] five.*

209 D *No. Ah, the difference about what is happening? Well, we can't control it.*

210 S *This is interesting. This could be the whole "catch" of our hypothesis!*

Shay's emergent role as a researcher was strengthened by the modeling task design approach. His motive to continually improve the model based on additional data led him to suggest a new conjecture. Using the AG graphing technique, he noticed surprising phenomena: the bird's motion might be nonlinear. Furthermore, he expected that this discovery was the key to understanding the bird's real motion [210]. While collecting more data to test this discovery, Shay expressed his main goal: "Should we provide a mathematical formula that really explains many stretch levels? … Because just for three levels of stretches the game isn't interesting" [212]. Thus, Shay's researcher role led him to interpolate, extrapolate and consider a general function that would capture the bird's movement in a fuller way. Shay however didn't abandon the gamer hat to motivate the need for such a formula in order to make the game more interesting. Dor the gamer on the other hand was not impressed by Shay's discovery, and referred to it as irrelevant. Shay and Dor continued to refine the model and concluded the activity with different reports to the manufacturer, which reflected their different perspectives and goals related to their emergent roles.

DISCUSSION

The examples above provide insights on students' emerging roles in a statistics task. Shay and Dor spontaneously took two roles: The *researcher* and the *gamer*. The researcher role grew from the integration of the EDA and the AG design approaches and the students' fluency in norms and dispositions of data exploration. The AG approach strengthened the researcher's role, by adding the need to reason iteratively about data and graphs to refine conjectures and raise questions about the next inquiry step. The gamer role grew clearly from the gaming approach. The boys designed a game that was not only based on the real context, but was also exciting and interesting.

While these three approaches played a role in the emergence of students' different goals and roles, the modeling approach had remarkably created opportunities for conflicts between the roles. The requirement to draw a model of the bird's movement forced the boys to negotiate their beliefs, roles and points of view. Each of them however seemed to hold both roles and these conflicts made them choose the "leading role" at times. Both roles played an important part in the development of students' reasoning about uncertainty. The researcher role encouraged the students to acknowledge uncertainty and define it as accurately as possible, search for patterns and use scientific and statistical terms and methods. The gamer role led in some cases to deliberately ignore uncertainty, but in other cases created the need to better control uncertainty and search for simpler abstractions of it. The conflict between the two roles and the points of switching the leading role resulted in the refinement of students' goals and consequently their conjectures and models.

This concise account is far from exhausting this complex learning phenomenon. The integration of the four task design approaches seems to be fruitful, but further research is essential to examine the connections between the task design and students' spontaneous roles, the role of various activities in the emergence of roles, other types of roles that might emerge and their role in developing students' statistical reasoning. Other factors that might influence role-taking should be considered, such as working in a larger group than a pair, or the task design as a part of a larger learning trajectory. In any case, it seems that this new line of research can advance our ongoing vision and efforts to understand and improve the learning of statistics.

ENDNOTES

[1] This study was supported by the British Academy Small Research Grant Scheme (SG112288). The views expressed in this paper do not necessarily reflect the views or policy of the British Academy.

[2] The translation from Hebrew to English focused on preserving the authenticity of the original utterances, sometimes at the expense of correct English phrasing.

[3] The numbers in square brackets refer to the number of the quote as it appears on the transcript of the episode.

[4] Insertions in square brackets represent our attempt to complete what students mean given the context and their gestures.

REFERENCES

Ainley, J., Aridor, K., Ben-Zvi, D., Manor, H., & Pratt, D. (2013). Children's expressions of uncertainty in statistical modeling. In J. Garfield (Ed.), *Proceedings of the Eighth International Research Forum on Statistical Reasoning, Thinking, and Literacy* (SRTL-8). Minneapolis, MN, USA: University of Minnesota.

Ainley, J., Nardi, E., & Pratt, D. (2000). Towards the construction of meanings for trend in active graphing. *The International Journal of Computers for Mathematical Learning*, *5*(2), 85-114.

Ainley, J., Pratt, D., & Nardi, E. (2001). Normalising: Children's activity to construct meanings for trend. *Educational Studies in Mathematics*, *45*, 131-146.

Biehler, R. (1990). Changing conceptions of statistics: A problem area for teacher education. In A. Hawkins (Ed.), *Proceedings of the International Statistical Institute Round Table Conference* (pp. 20-38). Voorburg, The Netherlands: International Statistical Institute.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer.

Gil, E., & Ben-Zvi, D. (2011). Explanations and context in the emergence of students' informal inferential reasoning. *Mathematical Thinking and Learning*, *13*(1&2), 87-108.

Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, NY: Doubleday.

Hare, A. P. (1994). Types of roles in small groups: A bit of history and a current perspective. *Small Group Research*, *25*, 443-448.

Jahnke, I. (2010). Dynamics of social roles in a knowledge management community. *Computers in Human Behavior*, 26, 533-546.

Konold, C., Harradine, A., & Kazak, S. (2007). Understanding distributions by modeling them. *International Journal of Computers for Mathematical Learning*, *12*(3), 217-230.

Manor, H., Ben-Zvi, D., & Aridor, K. (2013). Students' emergent reasoning about uncertainty while building informal confidence intervals in an "integrated approach." In J. Garfield (Ed.), *Proceedings of the Eighth International Research Forum on Statistical Reasoning, Thinking, and Literacy* (SRTL-8). Minneapolis, MN, USA: University of Minnesota.

Moore, D. S. (1997). New pedagogy and new content: the case of statistics. *International Statistical Review*, *65*, 123-137.

Morris, R., Hadwin, A. F., Gress, C. L. Z., Miller, M., Fior, M., Church, H., & Winne, P. H. (2010). Designing roles, scripts, and prompts to support CSCL in Study. *Computers in Human Behavior*, *26*, 815-824.

Mudrack, P., & Farrell, G. (1995). An examination of functional role behavior and its consequences for individuals in group settings. *Small Group Research*, *26*, 542-571.

Pfannkuch, M., Rubick, A., & Yoon, C. (2002). Statistical thinking: An exploration into students' variation-type thinking. *New England Mathematics Journal*, *34*(2), 82-98.

Salen, K. (2007). Gaming literacies: A game design study in action. *Journal of Educational Multimedia and Hypermedia*, *16*(3), 301-322.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. NY; Basic Books.

Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R.S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (Vol. 2, 6th ed., pp. 464-510). Hoboken, NJ: Wiley.

Slavin, R. E. (1995). A model of effective instruction. *Educational Forum*, *59*, 166-176.

Strijbos, J., & De Laat, M. F. (2010). Developing the role concept for computer-supported collaborative learning: an explorative synthesis. *Computers in Human Behavior 26*(4), 495-505.

Tukey, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.