UNDERSTANDING BASIC DEMOGRAPHIC TRENDS:
CONNECTING TABLE READING, TASK DESIGN, AND CONTEXT

Iddo Gal & Anna Trostianitser
Department of Human Services, University of Haifa, Israel
<iddo@research.haifa.ac.il>

The paper aims to inform the conceptualization of teaching/learning goals and instructional design related to understanding multivariate statistics that describe trends in society. The paper presents a conceptual analysis leading to identification of a gap in current analyses and instructional sequences focused on interpretation of information in tables [and graphs]. The paper advocates for the need to develop students' ability to reflect about the factors that cause or are related to observed trends, and about their societal implications, and highlights the need to develop task design principles that can promote such understanding. We then present an empirical demonstration of a possible class sequence, and results from an exploratory class activity with 43 students whose responses were content-analysed. Based on these results and the prior analysis, we reflect on benefits in terms of increased motivation and relevance of statistics instruction, and call for the use of task design principles that directly connect data, statistics, and class activities to the actual societal context in which students have to operate as adults, in order to enhance students' statistical literacy and understanding of statistics about society.

BACKGROUND

An important goal of schooling is to prepare school graduates to take part in societal processes and become engaged and empowered citizens. In the area of statistics education, an argument repeated by many scholars (e.g., Gal, 2002; Ridgway, 2015) is that the learning of statistics can help students to become informed citizens who can engage statistical messages with a critical perspective and can not only interpret statistical messages (e.g., in the media) but also relate the soundness of statistical arguments. However, bringing students to that stage is an uphill battle, given many challenges involved in statistics education (Batanero, Burrill, & Reading, 2011).

Several challenges exist in statistics education. According to Ograjensek & Gal (2016), a critical challenge involves issues of students' acceptance of *purpose*, i.e., understanding the "Need to Know", i.e., the purpose for the whole exercise of interpreting data, and the value of good data, within the broader environment or societal context which statistics aims to inform. Another is the *motivational* aspect of how to make the technical and conceptual knowledge learned within a statistics class appear relevant and interesting to students. Both of these challenges, regarding purpose and motivation, become especially pressing when trying to connect students with official statistics which describes important societal phenomena.

This paper focuses on a specific aspect in this regard, that of engaging students with statistics about trends in society, e.g., regarding demographic changes, trends in crime or educational achievement or economic variables, to give just a few examples. Our work in this area is part of the ProCivicStat project funded by the European Commission's ERASMUS+ program. ProCivicStat argues that data about civic topics are usually produced by official statistics agencies and pertain to dynamic phenomena that change over time and require attention to multiple subgroups, hence are multivariate in nature. Further details about the logic underlying ProCivicStat are provided by Engel, Gal, & Ridgway (2016), and further details about civic statistics tasks are given in Gal, Ridgway & Nicholson (2016) and by other publications on the ProCivicStat website at: http://community.dur.ac.uk/procivic.stat.

Specifically, data products from official statistics agencies that aim for the general public, such as press releases, often involve tables and graphs combined with descriptive tests. A literature search we have conducted shows that very few studies focus on issues associated with teaching and learning of such data products that are in the public sphere. While statistics producers create data products that contain essential information regarding trends and patterns in society, they are normally not included in traditional approaches to teaching statistics, which focus on statistical procedures and not on the context from which the data emerges and on their societal implications.

Given the above, a new challenge emerges, regarding how to create motivated students who are willing and ready to engage products of official statistics agencies and interpret them in a meaningful way, and how to help them see the value of official statistics in exploring trends in society. The present study focuses on a subset of this area, related to students' work on students work on data about society that is presented in tabular form, given its ubiquity.

Numerous scholars have examined issues associated with reading and interpreting information in tables and graphs, given their core role in statistics education (Sharma, 2013). The emphasis has usually been on cognitive operations associated with harvesting quantitative information and interpret the patterns in the data. Curcio's well known (1987) framework has discussed three stages or steps, i.e., reading the data, reading between the data, and reading beyond the data. Based on these foundations, Friel, Curcio and Bright (2001) have outlined six cognitive processes: reading, describing, interpreting, analysing, predicting and extrapolating data.

Building from these and other sources, Kemp and Kisanne (2010) offered a five-step framework that can serve as a guide for class discussions: 1: Getting started; 2: WHAT do the numbers mean? 3: HOW do they change or differ? 4: WHERE are the differences or relationships? and 5: WHY do they change? Recently, Prodromou (2015) has demonstrated how Curcio's original conceptualization can be applied as part of group-work of students interpreting large 2-way tables comparing population trends (demographic changes, emigration and immigration, etc) across European countries.

The frameworks described above have contributed to an important body of knowledge regarding the thinking processes that students should engage with when learning to read tables and graphs, and possible building blocks for an instructional sequence, e.g., how to create questions that walk students from low-level technical aspects to broad interpretative issues. However, the emphasis in the above and other studies has been on the statistical or mathematical aspects of reading tables and graphs. We have little knowledge about how students connect between data and the broader societal context, and what external knowledge they bring to bear on societal data they have to analyse in class.

The present paper aims to further explore how to address questions of purpose and motivation when using data or products of official statistics agencies that involve statistics about society. Specifically, we see a need to extend the five-step framework introduced by Kemp & Kisanne (2010) by adding a new step that focuses on the actual [societal] meaning or implications of the data. We believe that it is essential to understand how students reason both about factors that affect observed trends in data, and what might be the implications of the observed data, as only in this way can students engage the meaning of the data and their relevance to societal issues.

We also see a need to change question posing and task design so as to structure the analysis of tabular and graphical data and make sure that students focus on changes in the value of key variables. Issues of task design have received much attention in the literature on mathematics and science education, yet less attention in the area of statistics education. Gal (1998) has discussed the design of tasks that can elicit open-ended opinions of value to learning principles of table analysis. Ainley, Gould & Pratt (2015) have argued for the use of tasks that have two characteristics, a clear purpose and present an engaging challenge for the learners, whether or not they refer to a real-world application. However, the nature of civic statistics is different than the data or findings often used in introductory statistics, and any question about them should clearly relate to the societal context, and reflect their characteristics outlined earlier, e.g., in terms of being multivariate, aggregated, text-based, and the like.

With the above needs and directions in mind, we next describe a demonstration that takes students through the five steps described by Kemp & Kisanne (2010) but adds a sixth step as explained above, and changes some of the way earlier questions are posed, and implements some of the task design principles outlined above. The demonstration is conducted with students who are all studying a specific discipline, to enable the couching of questions in a context that is meaningful from a societal perspective. To be sure, the focus of the demonstration is not on how well students are able to interpret information in a given table [or graph], although this of course is of importance, but on students' reactions and thoughts regarding the new aspects of the tasks, which are linked to motivation and of purpose regarding statistics about society.

AN EMPIRICAL DEMONTSRATION

Approach. Display 1 shows a task that was given to 43 students in a first-year course "Introduction to Human Services" that the first author is teaching in a Department focused on preparing students to manage service operations in organizations in the public and private sectors. The class was multicultural with about 65% Jewish and 35% from the Arab sector (a broad category comprised of Muslims, Christians, and other cultural groups). The average age of participants is 22 years. 80% were studying in parallel in a standard introductory statistics class.

As part of this course, students hear a lecture about diversity of clients (e.g., in terms of demographics and other variables). Within this lecture, students were given the task shown in Display 1. The task was presented as a management simulation related to understanding client diversity, i.e., students had to operate in a managerial context in which they can be a real actor, not a passive or disinterested observer who just analyses data. In addition to answering the five core questions shown in Table 1, students also had to rate their interest in the task, its perceived difficulty, whether they studied statistics, and provide other background information.

Display 1. Sample task - Table reading with opinion elicitation

Central Bureau of Statistics: Projections 2005 - 2030
Note: Figures are in Thousands

Total (000)	Arabs		Jews		*2005*
	%	Number (000)	%	Number (000)	Age
	59%	800	41%	2,284	0-24
	38%	512	48%	2,659	25-64
	3%	44	11%	629	+65
6,928 =	100%	1,356	100%	5,572	TOTAL

Total (000)	Arabs		Jews		*2030*
	%	Number (000)	%	Number (000)	Age
	47%	1,115	35%	2,737	0-24
	46%	1,082	51%	4,004	25-64
	7%	165	14%	1,129	+65
10,232=	100%	2,362	100%	7,870	TOTAL

Managerial simulation

You are a new, motivated manager at organization X.

You received from your senior management the following table, which is based on new data from the Central Bureau of Statistics.

The table shows projections regarding demographic trends among certain social and age groups.

Your manager asked to get your personal evaluation regarding the following key questions.

Q1. What key trends or changes you see between the social and age groups, for 2005?
Q2. What key trends or changes you see between the social and age groups, for 2030?
Q3. What key trends or changes do you see between 2005 and 2030?
Q4. What reasons or factors may explain the projected changes between 2005 and 2030, regarding these social and age groups? please list key ones.
Q5. Given these trends or changes, what are the implications or impact for long-range planning or needed managerial decisions in [organization]? Why?

The data: The 3-way data table we used is based on population projections published by the Israeli Bureau of Statistics; it is comprised only of counts (in thousands) and percentages, which are typical in demographic tables. Thus, the data table employs seemingly simple statistics that should be accessible to all students at the undergraduate or high-school level even without learning any statistics. Yet, the table overall contain many elements, and is richer than typical 2-way tables employed in learning cross-tabulation. The table has a basic 2x3x2 design (i.e., social group, age, year), which is needed to present a basic multivariate problem. The data relate to anticipated changes in the demographic composition of the two largest subgroups in Israel, Jewish and Arab, hence provide a rich social context and couched in a reality known to all students.

Question posing: The five key questions listed in Display 1 are organized in a somewhat different way compared to the sequence proposed by Kemp and Kisanne (2010), which does not impose any structure on the actual process of [visual] analysis of the data. Questions Q1 to Q3 cover the first four steps in the Kemp & Kisanne (2010) model described above. The first three questions aim to break down for students the table analysis into steps, i.e., first asking about trends or changes in 2005 (top half of the table), then in 2030 (bottom half) and finally across the two timeframes (whole table). This approach to question posing was employed in order to ensure that students pay attention to the key variables along which the data are organized, and can gradually build a picture of trends and changes within the data, from bivariate to multivariate.

Questions 4 and 5 are the critical ones from the point of view of purpose and motivation, and reflect our quest for task design principles that directly connect learned with the context. Q4 seemingly addresses the same goal as Step 5 proposed by Kemp & Kisanne (WHY do they change). However, the phrasing we used is more structured and tailored to the specific context. Q5 adds a new layer and requires that students think about the actual societal or organizational *implications* of the data. Together, Q4 and Q5 take the student from a mere analysis of how numbers change with the cells of a table, to engagement with a real-world context and to thinking about the social meaning or factors affecting a socially-loaded issue.

FINDINGS

Due to space limits we present a selected summary of students' responses. Our main purpose is to illustrate the range of responses received and through them the educational potential of such a task. Before starting, we note that the task overall was allocated 30 minutes of class time; most students reached the end of the questionnaire and answered all five core questions in writing, but 6 of the 43 (about 14%) did not answer Q4 and Q5 and later explained they were still thinking about Q3, or contemplating how to phrase their answers Q4 but without writing anything about it, or just exhausted.

Q1 asked students to describe differences or trends they see in 2005 in relation to population and age groups. Content analysis identified 17 different themes in the responses of our 43 students, The most frequent themes were: differences in relative and absolute size of the young or old age groups across the two social groups, which age group is largest in each social group (either in terms of absolute numbers or by reference to proportion or ratio of young to old or old to young), differences in life expectancy among the two social groups, and comparison of the overall sizes between populations (i.e. ,more Jews than Arabs).

Content analysis of responses to Q2 generated 17 themes of quite a similar nature, with a slightly higher tendency to relate to ratios or proportion of one group compared to another, e.g., relative size of elderly group in Jewish and Arab sectors, or comparison of the relative size of the young group vs. old group within each social group size.

Q3 required students to integrate information from all parts of the data table across the 2005 and 2030 timeframes. Content analysis identified 20 themes, with the following five themes being the most common (between 44% to 14% of responses):

- life expectancy changes (44%): *"The number of people in the age group 65+ is growing in both population groups."* or ""*Another increase in life expectancy is expected, number of elderly people in the country in 2030 is increasing in Jewish and Arab population, compared to data in 2005."*

- General population growth (40%) *"It seems that the country's population (Jews and Arabs) will increase 1.5 times over the 25 years."* or *"The number of people in each sector and each age group has increased over 25 years."*

- Decline in young population (26%): *"Birth rate is falling despite the increase in the number of people. The percentage of those aged 0-24 decreases from 2005 to 2030 by 6% in Jews and for 12% in Arabs."*

- Increase in adult (middle) population (21%): *"Amount of people in the age group 25-64 nearly doubled in two population groups."*

- Growth of Arab sector (14%): *"Amount of Arab people almost doubles in 2030."*

Taken together, the responses to Q1, Q2 and Q3 show that students are able to read the information in the tables and discuss it correctly in various ways, though there is much diversity in what they refer to, with some responses related to reading horizontally (e.g., comparing size of two social groups within the same age) while others reading vertically (e.g., comparing the relative size of two age groups within the same social group). The language of some of the responses is not specific enough to understand if students are referring to changes in the absolute numbers or thinking in proportional terms.

Q4 asked students to list reasons or factors that may explain the changes they observed between 2005 and 2030. Content analysis of the 37 responses to this question identified 24 different themes. Note that students identified in Q3 many possible changes between 2005 and 2030, hence the range of explanations is also wide and different explanations may refer to different features in the data that particular students observed. The five most frequent themes were:

- Advances in medical services (60%): *"Developments in the medical field enhance longevity with use of new effective drugs."*

- Increased life expectancy (40%) *"Life expectancy has increased – percentage of elderly population is increasing."*

- Diminished birth rate (31%): *"Difficulty in establishing large families / economic crises, the cost of living."* Or: *"Today, many families in both sectors have decided to reduce the scope of the family and raise fewer children, so there is a decrease in this age group."*

- Improved standard of living (17%): *"Life becomes more modern and convenient and it helps elderly people to love their life."* Or: *""Emphasis on quality of life as a sports, food and spirit will naturally increase life expectancy."*

- Changes in technology (17%)*: "Increase in an advanced technology and increase in a variety of databases."*

Finally, Q5 asked students to indicate possible *consequences* for long-term planning or managerial processes in light of the trends or changes displayed in the table between years 2005 and 2030. We have identified 27 themes in content analysis. Here are examples for common ones:

- Allocation of resources in favour of elderly clients (33%): *"Elderly people should be more supported, to give them more care than to children.", Or: "In healthcare, to focus on the treatment of the elderly than children because their percentage will increase."*

- Effects of overall population growth on service planning and matching supply and demand (33%): *"The entire population will increase, for example, as the local authority manager, should take care of the population, housing for young people, etc."* Or: *"More population will require more kindergartens, schools, food. We should start planning from now on when and where will be built new buildings, train teachers, make sure there would be professionals that can provide these services...start to think about the future now because it take time to achieve goals"*

- Arab population growth and integration (14%): *"Organizations should take care to better integrate the Arab sector because this population will significantly grow."*

- Matching technology (11%): *"Improve and promote services of banks and in storage of bank money and the welfare of adults."* Or: *"The number of people aged 25-64 and 65+ will increase, hence the systems should be adapted to their technical knowledge."*

DISCUSSION AND CONCLUSIONS
This paper contributes to statistics education in several ways. It articulates the need to revisit ways for teaching about official data about societal phenomena, and for using principles of task design that can clearly connect students with the actual societal context. In doing so we address issues of motivation and sense of purpose (Gal & Ograjensek, 2016) in order to make sure that students not only are willing to engage with unfamiliar data but also are able to and interested in thinking about the meaning of data and articulating judgments about it. The paper demonstrates that it is possible to design activities that engage students with the analysis of multivariate data

about important social phenomena, in ways that are different than by using standard multivariate statistics (e.g., multiple regression), which cannot extract the kind of information that is gained from reflecting about descriptive displays.

It is useful to note that most students rated the demonstration task as difficult or challenging, yet 82% saw it as very interesting, and indeed kept working on it for close to 30 minutes. When asked afterwards about the task, verbal responses suggested that such a task is interesting because the data connects with real life, e.g.,: *"It is very interesting to me to see the difference between 2 periods and factors that cause them.*" Or: "*It is exciting to feel yourself a manager, try to think like manager and prepare for the future to come."* At the same time, a few students found some of the questions as demanding, and had trouble expressing their thoughts in clear language.

As was explained above, traditional approaches have focused on understanding the data (this is captured by Q1, Q2 and Q3 in our demonstration task) and what can be projected from it in *statistical* terms. However, this by itself reflects an internal view of statistics, and cannot by itself create reflective, statistically literate students. The design of instructional tasks needs to pay separate attention to the analysis of quantitative or statistical information (whether in tables, graphs, or other representations), and to their *examination in context*. How can contextualization be achieved? We argue that there is a need to go beyond the model proposed by Kemp & Kisanne (2010), as informative as it is, and pose questions both about causal factors or correlates and about consequences or implications of the trends in the data (Gal, 1998; Gal, 2002).

The need for students to address and discuss in class both data and context may seem questionable to those who believe that the goal of statistics education is to develop procedural mastery or focus on the understanding of the mathematical information contained in various data representations, or who feel uncomfortable to discuss in a statistics class anything that is not directly explicit in the given numerical data. Indeed, in our demonstration task, students' responses to Q4 and Q5 cannot be based on the numbers given in the data table, but on their general world knowledge, professional training, or common sense. However, we point out that the distinction between findings (i.e., statistical trends in the data) and their interpretation and implications is a standard practice in all scientific reporting. Such a distinction is formally forced upon all scientists by the expectation that a chapter about "findings" appears as a separate entity from a chapter that discusses conclusions, implications, limitations and caveats, and recommendations.

Our analysis of students' responses, while exploratory in nature, suggest that the interpretation of multi-way tables regarding a socially important phenomena is a complex undertaking for many students ─ yet is feasible in a heterogeneous class and has much educational potential as it can give rise to many different observations and informed conclusions, which are an important end-product of any data analysis process (Gal 1998). Thinking about the meaning of the data that is "enveloped" in questions about causes and correlates on the one hand, and about implications on the other, requires students to make direct connections to the social context from which the data emerge. Indeed, students bring rich interpretations based on multiple perspective, as reflected in the over 20 separate themes of responses to Q4 and Q5. This is what makes the task more interesting - yet more complicated from a teacher's point of view, as there are many "sensible" opinions or responses (not necessarily right or wrong in the usual sense). The pilot data illustrate the importance of question-posing and opinion-eliciting questions (see Gal, 1998) as factors that can increase the relevance and value of a statistical activity using socially relevant data, and help to motivate students to engage challenging parts of the task.

Students in the 21st century graduate into a world with more data of a more open nature (Ridgway, 2015). This requires that students take a more active role and are able to generate informed opinions and feel comfortable to connect between data and context, and know how to build good arguments that are based on data, though also on general world knowledge. We believe that reaching such goals calls for the need to revisit how we design instructional tasks that more directly connect data, statistics, and class activities to the actual societal context in which students have to operate as adults.

REFERENCES

Ainley, J., Gould, R., & Pratt, D. (2015). Learning to reason from samples: commentary from the perspectives of task design and the emergence of "big data". *Educational Studies in Mathematics*, *88*(3), 405-412.

Batanero, C, Burrill, G., & Reading, C. (2011). *Teaching Statistics in School Mathematics – Challenges for Teaching and Teacher Education.* (18th Study of the International Commission on Mathematical Instruction). Heidelberg: Springer.

Curcio, F. R. (1989). *Developing Graph Comprehension: Elementary and Middle School Activities*. Reston, VA: National Council of Teachers of Mathematics,

Friel, S. N. Curcio, F. R. & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional applications. *Journal for Research in Mathematics Education, 32*(2), 124-158.

Gal, I. (1998). Assessing statistical knowledge as it relates to students' interpretation of data. In S. Lajoie (Ed.), *Reflections on statistics: Learning, teaching, and assessment in grades K-12* (pp. 275-295). Mahwah, NJ: Lawrence Erlbaum.

Gal, I. (2002). Adult statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*(1), 1-25.

Gal, I. & Ograjensek, I. (2016). More on enhancing statistics education with qualitative ideas: A Rejoinder. *International Statistical Review,* 84(2), 202–209.

Gal, I., Ridgway, J., & Nicholson, J. R. (2016). Exploration of skills and conceptual knowledge needed for understanding statistics about society: A workshop. In J. Engel (Ed.) Proceedings, IASE 2016 Roundtable Berlin. Retrieved: http://iase-web.org/Conference_Proceedings.php.

Prodromou, T. (2015). Students' Emerging Reasoning about Data Tables of Large-Scale Data. *International Journal of Statistics and Probability*, *4*(3), 181-197.

Kemp, M., & Kissane, B. (2010). A five step framework for interpreting tables and graphs in their contexts. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia.* Online: http://iase-web.org/documents/papers/icots8/ICOTS8_7G3_KEMP.pdf

Ograjensek, I. & Gal, I. (2016). Enhancing statistics education by including qualitative research. *International Statistical Revie*w, 84, 2, 165–178. DOI: 10.1111/insr.12158

Ridgway, J. (2015 ahead-of-print). Implications of the Data Revolution for Statistics Education. *International Statistical Review*. doi:10.1111/insr.12110

Engel, J., Gal, I. & Ridgway, J. (2016). Mathematical literacy and citizen engagement: The role of Civic Statistics. Paper presented at the 13th International Conference on Mathematics Education, July 2016, Hamburg, Germany.

Sharma, S. (2013). Assessıng Students' Understandıng of Tables and Graphs: Implıcatıons for Teachıng and Research. *International Journal of Educational Research and Technology, 4,* 51-70. Online: www.soeagra.com/ijert/ijert.htm