CIVIC STATISTICS AND THE PREPARATION OF FUTURE SECONDARY SCHOOL MATHEMATICS TEACHERS

Achim Schiller[1] and Joachim Engel[1]
[1]Ludwigsburg University of Education, Germany
schiller01@ph-ludwigsburg.de, engel@ph-ludwigsburg.de

Statistics education in Germany at secondary and tertiary level is lagging behind the demands for engaged citizenship in the information age. We detected students' deficiencies in statistical literacy especially in regard to understanding multivariate phenomena which are common in data about society. Thus we started to change the way statistics is taught by introducing elements of what we call civic statistics *into the curriculum for students preparing to be secondary school mathematics teachers. Focusing on critical questioning and critical thinking one element is dealing with the understanding and interpretation of short data-based statements.*

INTRODUCTION

The availability of data has tremendously increased. The data deluge is changing the demands for engaged and informed citizenship in times of free access to data about social topics like income inequality, access to health services or migration through the internet, national and international statistics offices or NGOs. Understanding this kind of data requires different knowledge and dispositions than the current statistics education represents (Ridgway et al., 2013).

There are new dimensions for civic participation in public decision processes - to become a responsible citizen capable of critical and evidence based judgment about social topics as basis for committed involvement, the ability to understand and to interpret these data is indispensable (Engel, 2014). To understand these data citizens need specific statistical knowledge that is not part of the current curriculum, in addition to a cognitive and dispositional base that includes (but is not limited to) context knowledge and critical thinking skills (Gal, 2002).

Current statistics curricula in schools and colleges as well as in textbooks focus too much on formal statistical procedures as well as on uni- or bivariate small datasets, rather than on the knowledge and attitudes required for understanding patterns and changes in social phenomena. Most social phenomena are multivariate, presented by complex graphs or tables, have a laborious operationalization of variables, confounding variables and have correlations as well as multiple underlying causal factors.

Besides elements of descriptive statistics and basic concepts of exploratory data analysis needed for understanding the actual information, critical questioning as a knowledge element of statistical literacy (Gal, 2002) and "critical thinking as general thinking skills and strategies" (Aizikovitsh-Udi & Kuntze, 2014) are of equal importance.

TEACHING MODULES ABOUT CIVIC STATISTICS

In view of the needs for engaged citizenship we propose a new orientation in statistic education. Statistically literate citizen need the capacity to develop the knowledge, skills and attitudes to understand and communicate about relevant multivariate data. The project *Pro Civic Stat*, funded by the European Commission as strategic partnership between six universities, aims to develop new teaching modules for statistics instruction that contribute to young people's ability to be aware of and understand quantitative evidence and statistics about key social phenomena that permeate society, their social and private life. Selective information may develop into a stumbling block on the way to an individual's independent and informed interpretation. Therefore, it is important to know which information is necessary for a suitable valuation.

One of the goals of *Pro Civic Stat* is to design a complete course *Civic Statistics* comprising a set of self-sufficient modules which may also be integrated as smaller units into other (statistics) courses. Informed citizen as data-consumer (cf. Gal 2002, Rumsey 2002) need to be able to understand, verify, interpret, and critically evaluate data-related arguments and various

In: J. Engel (Ed.), *Promoting understanding of statistics about society.*
Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE), July 2016, Berlin, Germany. iase-web.org/Conference_Proceedings.php

representations (cf. Engel, 2014, Ridgway, 2015), understand concepts that describe social developments (cf. Engel, 2013) and differentiate between information and opinion.

THE CQS-MODULE

The content of the CQS-module is attributed to statistical and contextual knowledge and is supported by *critical questioning of data-based statements* (CQS). CQS combines **critical questions** with **critical thinking.** Critical questions relate to methods of data collection, the operationalization of variables or indices, the measurement process, the use of location and spread parameters, as well as graphical and other numerical representations. It also implies an awareness of possible confounding variables and the common fallacy of confusing correlation with causation. Critical thinking relates to the quality and provenance of data, the contextual knowledge, possible logical conclusions and the politics of data collection, presentation, interpretation or utilization.

With CQS we want to improve students' awareness of relevant information for a suitable assessment of data based statements. Furthermore we aim to improve their attitude towards statistics regarding its relevance, possibilities and implications for society. This module has already been field-tested twice (Feb and May 2016) in small courses for future secondary school mathematics teachers. The trials indicate that students have difficulties with CQS, because they barely ever were taught to ask critical questions at all about data or data-based statements. For example, they were not aware of the importance and intricacies of the data collection, the operationalization of the variables or of the selection of specific statistical summaries or methods.

The structure and content of the CQS-Module

We designed a CQS-module with the focus on critical questions about essential statistical information and context knowledge for an appropriate assessment of data-based statements. The CQS-module is designed to give students an overview of categories that are important to know about for a suitable assessment. The structure of the module is based on the Statistical Reasoning Learning Environment (SRLE) developed by Garfield and Ben-Zvi (2009). With the focus on developing central statistical ideas, using real and motivating data sets, as well as classroom activities, integrating appropriate technological tools, promoting classroom discourse and using assessment to monitor their development in statistical learning.

Commonly, the first contact with socially relevant topics occurs through headlines followed by a journalistic product in newspapers, the Internet or newscasts. It is impossible to communicate the total information contained in scientific data or a study to be presented to a general public. Much information has to be condensed for the sake of understanding and conciseness (See Figure 1).

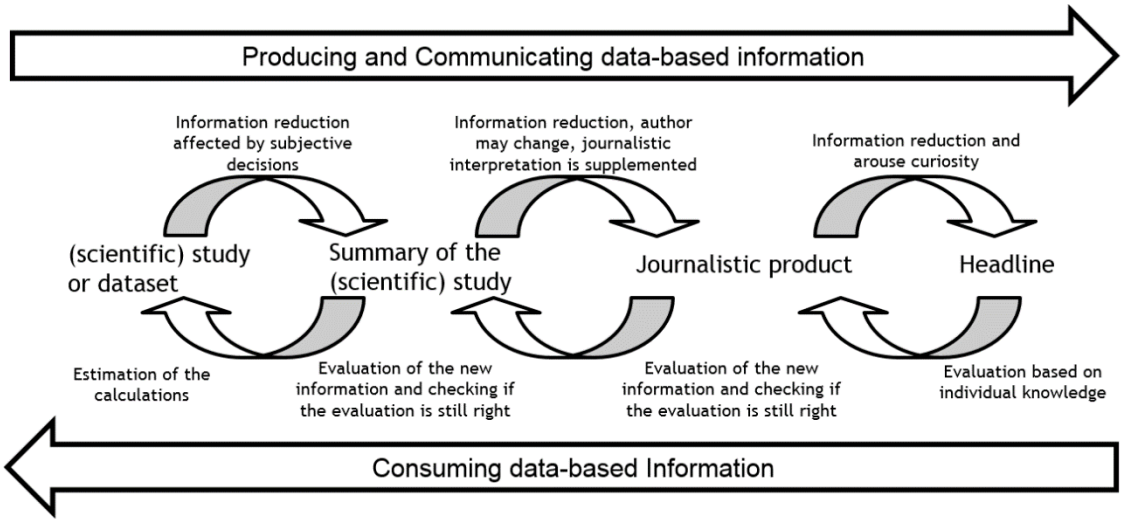

Figure 1: Order of Producing and Consuming data-based Information

The figure shows the path of producing and communicating data-based information from a (scientific) study to a headline. At each step there is a reduction of information that is influenced by

more or less subjective judgment. In addition, the reduction of information can be intensified by a change of the authors or other journalistic decisions. From the (scientific) paper to the journalistic product (subjective) opinions are usually added. The decision what is written in the headline is again subjective and usually orientated towards the clientele.

The structure of the CQS-module is leaned on the process of consuming information (See Figure 1). The figure shows the path of consuming data-based statements from the headline to the (scientific) study. The data consumer is usually addressed by a headline. The headlines are reduced to the supposedly most important message or summarized to a few numbers. The consumer evaluates the headline based on his individual knowledge and detects possible information deficits. These deficits are tried to compensate with additional information within the journalistic product. This information is usually very selective, based on an assessment of relevance by the media or journalists. The journalistic product is examined whether the necessary information for a critical and evidence based judgment is available. In addition, the consumer should differentiate between the original information of the producer and the opinion of the communicator. Increasingly, Internet links for the original data or the study results are given in the journalistic products. So the consumer has easy access to the sources of the communicator, which he can then compare with the journalistic product.

Our module started with four selected headlines from current news articles presented as single sentences of a data-based statement (e.g., *The humanity has never been so fat, for the first time the amount of obese people outnumbers the underweight*). The students were encouraged to rate these statements and to ask questions about which additional information may be needed for a well-founded rating of the given statements (e.g., "how big was the sample?", "was the sample representative?"). Discussing about these questions and ratings in class gives a first overview of what might be important to critically assess these statements. The diversity of evaluations among the students served as eye-opener to realise the legitimacy of different interpretation based on the same information.

Passing out the original articles as additional information about these headlines (e.g., http://www.zeit.de/wissen/gesundheit/2016-03/uebergewicht-adipositas-ernaehrung-bmi-entwicklung), enabled the students to reassess their initial evaluation, to answer some of their first questions or raise new ones. Discussing possible changes of the questions and ratings of the associated articles illustrates different foci of questions and ratings as well as different opinions based on the same information. The next step for the students was to classify their questions into self-imposed categories. In the final part the students were given different additional references. For example, links to homepages with interactive and dynamic visualization tools (See Figure 2) or summaries of (scientific) studies.

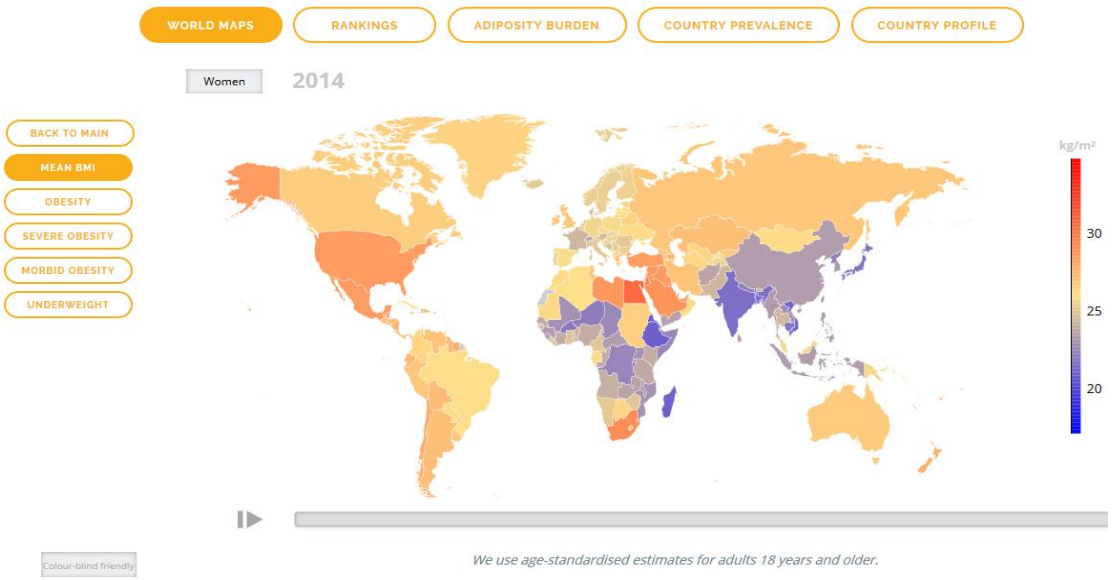

Figure 2: Evolution of women BMI over time
http://www.ncdrisc.org/map-mean-bmi.html

At the end the students arranged all their questions about the additional information may be needed for a well-founded rating of the given statements into categories that are important for a critical evidence-based judgment. The five categories found by the students during the plenary discussion reflect the fundamental statistical ideas of Burrill and Biehler (2011). (1) Data collection, (2) operationalization of variables and (3) context knowledge which Ridgway, Nicholson and McCusker subsumed under *Data* (2011). (4) Graphical or other representation discussed by Gonzalez, Espinel & Ainley (2011). (5) Statistical calculations with their strengths and weaknesses. These five categories have illustrated for the students that journalistic products have often too few information for a critical evidence-based judgment.

The final task of the CQS-module for the students was to evaluate and to interpret a given headline (e.g., *In the past not everything was cheaper. On average, life today is not more expensive than in 1991, http://www.zeit.de/wirtschaft/2014-05/inflation-loehne-deutschland*) without further information. Instead they were encouraged to research for additional information by themselves for a critical and evidence based judgment based on the five concepts of CQS.

At the beginning of CQS and the initial discussions about the given data-based statements, the students had difficulties to ask critical questions about such statements. Some questions about the representativeness of the sample, the type of data collection, the definition of the variables or inquiring about the story behind the data were achieved by slight impulses from the instructor. For example, it was quite difficult to elicit from students, how the variables were measured or if the conclusion were unambiguous. The first discussions suggest the assumption that these students didn't scrutinize the plausibility of such data-based statements.

The Goals of the CQS-Module

The first of three goals is to teach which information about data-based statements is important to know about for a suitable informed interpretation. Therefore it is important to know how the data are collected, how the relevant variables were operationalized, which statistical calculations are carried out, which scope of interpretation is given by the presentation of the results and in which context the study is embedded.

The second goal is to raise awareness and build up understanding that every step of producing, communicating and consuming information is influenced by subjective judgments, personal decisions or individual knowledge (see Figure 1) which may result in different interpretations. The choice of the type of data collection, the operationalization of the variables and the selection of statistical calculations are decisions based on many different criteria by the researchers. The presentation of the results as well as their interpretation in journalistic products reflect the perspective of the journalist as data communicator. The obtained results as well as the context knowledge are important for the interpretation. Different approaches to a topic may lead to different conclusions.

The third goal is that students learn to distinguish between (objective) information and (subjective) opinion and know that even with the same information contradictory interpretations may arise.

Implementation

We implemented the CQS-module in an optional small course for future secondary school mathematics teachers. In addition to the CQS-module the students generally discussed the role of statistics in society and which statistical basic knowledge is necessary or helpful in everyday life. For example which statistical parameters (e.g., mean, median or standard deviation) or indices (e.g., consumer price index or BMI) they encounter. Within the next sessions the students evaluated the new curriculum documents (state curriculum 2016 for mathematics for the state of Baden-Württemberg) whether it responds to the needs of future citizens and explicitly includes topics related to civic statistics. Followed by analysing tasks in the latest edition of widespread textbooks, the students checked if and how concepts of CQS occur. At the end of the course the students searched for newspaper articles with data related content and discussed how to use such articles (as good and bad examples) how to implement the topics of CQS in high school.

DISCUSSION AND IMPLICATION

In today's media citizens are confronted with data-based statements in compressed format on topics such as migration, health, economy or environment. As informed citizens they need skills to understand, evaluate and interpret these often compressed data-based statements. To inquire into more detail, citizens need skills and knowledge to deal with statistical results in context beyond those statistical literacy skills described by Gal (2002). The skills need to be extended to include the handling of multivariate (open) data-sets and data-based statements derived from them. A critical and evidence based judgment presupposes knowledge about the data collection, the operationalization of variables, the statistical calculations involved, representation and the context.

The testing of the CQS-module showed that at the beginning the students had problems to ask critical questions about data-based statements. They were not used to ask questions and they did not know which information they actually need for a correct valuation. Nor was the role clear of possible variation or subjective influences in the whole information producing, communicating and consuming process. The CQS-module is primarily based on journalistic products and the individual questions and evaluations of the students. With these elements important categories for understanding, evaluating and interpreting data-based statements could be identified. The used materials can be modified and adjusted to student's interest and current events. The next steps are further testing and evaluating of the teaching materials and developing a CQS-module for large university classes.

ACKNOWLEDGMENT:

The work reported in this paper was supported in part by ProCivicStat project, a strategic partnership of the Universities of Durham, Haifa, Ludwigsburg, Paderborn, Porto and Szeged, funded by the ERASMUS+ program of the European Commission. However the views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of the funding agency.

REFERENCES

Aizikovitsh-Udi, E. & Kuntze, S. (2014). Critical Thinking as an impact factor on statistical literacy – theoretical frameworks and results from an interview study. In Makar, K., de Sousa, B. & Gould, R. (Eds.). *Sustainability in statistics education: Proceedings of ICOTS9.* Voorburg, the Netherlands: Intern. Statistical Institute.

Burrill, G. & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In Batanero, C., Burrill, G. & Reading, C. (Eds.). *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study*, Springer, New York, S.57-69.

Engel, J. (2013). Statistics education and human rights monitoring. In S. Forbes and B. Phillips (Eds.), *Proceedings of the joint IASE/IAOS Satellite conference Macao*. http://iase-web.org/documents/papers/sat2013/IASE_IAOS_2013_Paper_2.5.1_Engel.pdf.

Engel, J. (2014). Open data, civil society and monitoring progress: challenges for statistics education. In Makar, K., de Sousa, B. & Gould, R. (Eds.). *Sustainability in statistics education: Proceedings of ICOTS9.* Voorburg, the Netherlands: Intern. Statistical Institute.

Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities, in International Statistical Review, 70, S.1-25, Netherlands.

Garfield, J. & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment, in Teaching Statistics 31, 3, S.72-77.

Gonzalez, T.M., Espinel, C.M. & Ainley, J. (2011). Teachers' graphical competence. In Batanero, C., Burrill, G. & Reading, C. (Eds.). *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study*, Springer, New York, S.187-198.

Ridgway, J., Nicholson, J. & McCusker, S. (2011). Developing statistical literacy in students and teachers. In Batanero, C., Burrill, G. & Reading, C. (Eds.). *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study*, Springer, New York, S.311-322.

Ridgway, J., Nicholson, J. & McCusker, S. (2013). Open data and the semantic web require a rethink on statistics teaching. Technology Innovations in Statistics Education, 7 (2).

Ridgway, J. (2015). Implications of the data revolution for statistics education. in International Statistical Review, http://onlinelibrary.wiley.com/doi/10.1111/insr.12110/full.

Rumsey, D.J. (2002). Discussion: Statistical Literacy: Implications for Teaching, Research, and Practice, in International Statistical Review, 70, S.32-36.