AN ANALYSIS OF MEDIA ITEMS ABOUT THE CORONAVIRUS PANDEMIC: NEW INSIGHTS FOR STATISTICAL LITERACY

Iddo Gal and Vincent Geiger
University of Haifa, Israel
Australian Catholic University, Brisbane, Australia
iddo@research.haifa.ac.il
vincent.geiger@acu.edu.au

Despite repeated calls for development of statistical literacy among citizens, in fact there is a dearth of recent systematic empirical research on the actual mathematical and statistical demands of mainstream media. Hence, this study aimed to develop a typology of the content of mathematical and statistical products and demands in the COVID-19 pandemic media, given the criticality of such information to citizens and societies alike. We conducted content-analysis of a purposive sample of over 300 media items from digital news sources based in four countries with different profiles. The analysis generated nine categories of new or enhanced types of knowledge and skill demands evident in the media items analyzed, such as regarding models and causality, data quality and strength of evidence, comparative thinking, literacy and language, official data sources, critical interpretation, and more. We discuss implications for current conceptual models and for instructional efforts focused on statistical literacy.

BACKGROUND

The mass media is considered the primary vehicle through which most citizens consume the news on key social and economic matters. Consequently, the nature of the knowledge and skills needed for critically interpreting media reports has been the focus of a number of lines of research enquiry, for example, on mathematical literacy, adult numeracy, and statistical literacy (Jablonka, 2003; Geiger, 2015). The COVID-19 (Coronavirus) pandemic, however, has led to the production of vast amounts of quantitative data and statistical information by various actors in the public and private sphere. This information and data has been utilized in the news media to communicate the status of the pandemic and to make predictions of its progress in a dramatic fashion, at times both contradictory and contentious, raising the need to rethink the capabilities needed for citizens to form judicious and balanced opinions about world affairs and navigate their own behavior.

This presentation examines the capabilities needed to understand the mathematical and statistical information included in the COVID-19 (Coronavirus) pandemic media, given the critical nature of such information for the well-being and safe practices of all citizens (e.g., Maass, 2019). During the Coronavirus pandemic, which is still ongoing in many parts of the world, citizens have been exposed to a wide range of media items regarding the progression and status of the pandemic and its many impacts. 'Media items' is used here as a general umbrella term encompassing diverse types of articles or communications in print-based newspapers and magazines, in digital media (e.g., news websites, TV news programs showing expert interviews or speeches of public officials), radio broadcasts, blogs, podcasts, and postings on social networks by official actors and private citizens.

Media items regarding the Coronavirus pandemic include a wide range of statistical and mathematical products ('StaMPs'), which encompass texts (written or spoken), numbers, and visual representations that communicate diverse types of statistical and mathematical ideas, for example, regarding quantities, rates, comparative figures, results of models and projections, among others. Citizens needed to make sense of the information and meaning conveyed by StaMPs in media items, in order to understand government (or their agencies) decisions about behavior in public contexts (e.g., in terms of social distancing, vaccination), as well as when making personal decisions in this regard. The extensive use of StaMPs in the pandemic media to inform as well as influence public opinion and behavior was fore-shadowed by Steen (1999), who argued that citizens must have the tools to critically evaluate arguments and opinions present in a 'data drenched' world.

Yet, despite the importance of citizens' ability to engage with statistical information about burning social issues (Gal, Nicholson & Ridgway, Forthcoming 2021), and developing theories about the building blocks of statistical literacy (e.g., Gal, 2002), there is a surprising absence of systematic empirical research on the actual numeracy and statistical literacy demands in mainstream media.

In: R Helenius, E Falck (Eds.), *Statistics Education in the Era of Data Science*
Proceedings of the Satellite conference of the International Association for Statistical Education (IASE),
Aug-Sept 2021, Online conference. iase-web.org/Conference_Proceedings.php

Hence, this paper contributes to new knowledge by presenting selected findings focused on the demands of media items related to the Coronavirus (COVID-19) pandemic. Through this approach, we aim to contribute to new knowledge regarding the mathematical and statistical capabilities that citizens need to comprehend STaMPs appearing within media items.

METHODOLOGY

We have adopted an exploratory approach through the content-analysis of over 300 'media items' (articles, videos, podcasts). These items were purposively selected from four leading news sources, listed in Table 1, which are aimed at general but diverse audiences in four countries (Australia, UK, USA, Israel) with heterogeneous demographic, geographical and economic characteristics, and different pandemic-related profiles of infection. The chosen media items appeared during the timeframe 15 March-15 June 2020, a period across which the pandemic progressed, reached its 1st-wave peak, then started to subside in many countries. About half of the media items were lead or section-lead articles, an indication of their perceived importance to readers. Our approach to content analysis involved inductive category formation, which Krippendorff (2004) suggests as appropriate when texts written about a common theme are realized by different authors. Content categories were developed through a constant comparison process, discussions between the researchers of incoming media reports, and refinement until the categories and their definitions stabilized.

Table 1. Characteristics of chosen news outlets

Source	Outlet characteristics
ABC, The Australian Broadcasting Corporation	A publicly funded media outlet with international reach. Viewed as a highly credible source of information about current affairs.
YNET, Yediot Ahronot (Israel)	The website of the largest circulation print-based newspaper in Israel (Hebrew-based), considered centrist and investigative.
CNN news network (International Edition), USA	A USA-based media outlet reporting on international events. Considered both credible and reliable, left-of-center perspective.
The Sun, United Kingdom	The website of a tabloid which is the highest-circulation newspaper in the UK. A populist media outlet, right-of-center perspective.

RESULTS AND ANALYSIS

Through our analysis of the chosen media items we have identified nine categories of elements or demands of STaMPs, summarized in Figure 1, that require mathematical and statistical knowledge and skill in order to understand the news communicated to the general public about the pandemic. (Of course, media items may include other content that is of much relevance to readers but does not involve StaMPs).

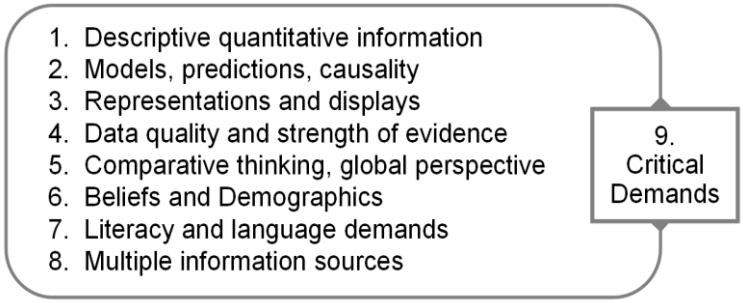

Figure 1. Nine categories of elements within the media items

We see categories 1-8 in Figure 1 as identifiably discrete, yet at times interrelated, since a single media item may encompass content from multiple categories. For example, results of modelling (Category 2) may be communicated in media items via spoken texts (Category 7) and/or diverse

representations (Category 3). Category 9, critical demands, however, is broader and can be relevant to or applied to any of the other eight categories when interrogating a media item. Most of these nine categories have seemingly been noted in the literature before, but not all within a single conceptual framework as presented here. All contain new or enhanced aspects that reflect salient demands in the pandemic-related news media.

These categories will be outlined as part of our presentation, and illustrated through selected media items. Due to space constraints, in this brief paper, we outline and provide illustrations related to only two categories, (a) models, predictions and causality, and (b) data quality and strength of evidence, which are at the heart of models of statistical literacy.

More on Category 2: Models, predictions, and causality

A model is a simplification and representation of reality that incorporates essential features of a real-world situation and the relationships between them. Statistical models are used for predicting the progression of many phenomena including the progress of epidemiological events. Models are complex tools, and their outputs must be interpreted with caution because of their reliance on assumptions and the quality and quantity of available data (e.g., Maass, et al., 2019).

Many key predictions and targets based on modelling have been reported in the pandemic-related media, such as the *rate of progression* of the pandemic over time in terms of, for example the number of infections, people in intensive care, or deaths. Modelling has also been used to predict the potential effect of various interventions (e.g., social distancing) on the progression of the pandemic, in order to avoid available medical resources being overwhelmed (i.e., how to '*flatten the curve*') or projections for when or under what circumstances *herd immunity* can be achieved, if at all. In addition, media items point to a range of seemingly simpler (though not simple) *indicators* which are models used not for purposes of prediction but first of all for purposes of data aggregation. Examples are 'positivity rate' 'reproduction rate' or R-rate, "rolling 7-day average", or "deaths per 100,000", none of which are included in the repertoire of ideas in regular statistics classes, let alone mathematics classes at the high-school level.

Models are thus used by governments, taking account of the advice of experts, to provide a basis for formulating policies aimed at controlling the pandemic and alleviating its social and economic impacts. Modelling results were pushed into mainstream public discourse via multiple media channels and digital platforms, at an unprecedented level, in order to convince populations to accept proposed or obligatory national responses. The public visibility of discussions about models and modelling, including the *vagueness or error* inherent in their predictions, was raised by well-publicized press briefings by heads of state (presidents, prime ministers), key public figures, and in statements by experts, in all news outlets included in our sample. This aspect of COVID-19 related reporting is illustrated Table 2, showing an excerpt from an ABC article.

Table 2. Models and vagueness in a media article (ABC, excerpt from Hayne, 2020)

Coronavirus 'nowcasting' modelling shows Australian case numbers continue to fall

...as health authorities continue to refine modelling of Australia's coronavirus curve, the margin of error in their findings is getting wider — but that's actually a good thing. Chief Medical Officer Brendan Murphy presented the latest coronavirus modelling on Friday afternoon, showing that Australia continues to stamp down the number of new infections being recorded each day.

Currently, there have been 6,675 confirmed cases of COVID-19 in Australia, of which more than 5,000 have recovered. There have been 78 deaths.

Margins of error are getting wider

The Government's "nowcasting" of the coronavirus situation in Australia aims to take stock of the coronavirus situation using the latest numbers. But that process gets less accurate if less data is fed into it. As case numbers continue to fall, the data being put into modelling is shrinking, making forecasts less precise. That being said, predictive modelling looking at the coming two weeks does suggest a further decline in numbers...

Discussion of models and the results of modelling were often intertwined with commentary on cause-and-effect relationships – a hallmark of scientific and statistical thinking. This was often seen in items which discussed causal influences and long-range *consequences* of the pandemic, such as reduced economic activity, unemployment, increases in domestic violence, or long-term health effects, each of which require policy interventions. Further, models, and their underlying assumptions or correlates, have been typically described in the media items via *text or verbally* (in the case of audio or video reports), only sporadically with the aid of graphs that show hypothetical or projected distributions. All the above means that understanding models and the impact of their predictions requires reading comprehension as much as mathematical and statistical facility, going well beyond the basic ideas that are normally taught about models (e.g., when learning about simple regression) in a typical introductory statistics class.

More on Category 4: Data quality and strength of evidence

Numerical data are typically presented within mathematics classrooms as complete, stable and accurate. Students learn that even *uncertainty* can be quantified via the laws of probability – usually in the abstract and with no real-world context attached to the computation. Notions of *error* (which may increase uncertainty) are usually covered in introductory statistics textbooks only regarding specific technical issues, such as error due to sampling, survey non-response, or computational methods. Accordingly, lists of "worry questions" have been proposed (e.g., Gal, 2002), describing topics about which readers of statistical messages should ask critical questions.

Yet our sample of media items about the pandemic covered many *additional* notions that go beyond 'error' in the technical sense as they also incorporated notions such as ambiguity, inaccuracy, variability, trustworthiness, or quality of evidence. These ideas were prominent in all media outlets analyzed and encompassed two elements (illustrative examples omitted due to space constraints).

- *Accuracy, reliability, or validity of descriptive figures, estimates and predictions*. Many media articles show that citizens should question basic data, indicators, or related public statements, because they may contain errors or be misleading. The media points to the need to look beyond national summaries to the underlying distributions and variability of subgroups. There were also instances of different predictions for the same phenomena (e.g., expected deaths) presumably all based on credible models.
- *Strength of evidence, and impact on policy and individual decisions.* Numerous media articles discussed 'evidence' in broad terms, not in reference to specific technical issue, suggesting that readers need to understand that quantitative information can vary in its *overall* comprehensiveness, acceptability or trustworthiness. There was discussion in media items about the willingness of certain social groups (or even public figures) to accept questionable evidence or questionable news (e.g., "fake news") that involved potentially misleading statistics, and the need for fact checking of data-related statements. The challenge facing governments in shaping policy based on inconclusive evidence was also discussed in some media items, as in debates regarding evidence supporting policy on the contentious issue of wearing face covering, as illustrated in the CNN excerpt in Table 3.

Table 3. Scotland recommends face coverings as cracks emerge in UK-wide approach to Coronavirus (CNN, excerpt from Rahim, 2020)

Scotland's First Minister said..."The evidence on the use of face coverings is limited, but there may be some benefit in wearing a facial covering when you leave the house and enter enclosed spaces."…. [In England] the Scientific Advisory Group for Emergencies (SAGE) have submitted evidence to ministers, and the government would announce a decision as soon as it was made, the Prime Minister's official spokesperson said,....Deputy Chief Scientific Adviser Angela McLean said SAGE had concluded there was "weak evidence of a small effect" in which face masks could prevent an infected person passing coronavirus on to someone else.

Overall, the media items portray the *tentative nature* of pandemic-related statistics and predictions. Texts suggest to their readers that raw data, analyses, and results can all be *revised* or *contested,* challenging traditional truth-laden views of 'scientific' statistics.

DISCUSSION

Our sample of media items is limited given the huge volume of information in the media about the pandemic, and for now the findings do not relate to media reports about later stages of the pandemic, including additional waves caused by COVID-19 variants, or vaccination programs and their effects. Yet, the sources used were sufficiently rich for results to contribute to new knowledge by pointing to a range of demands in the media that have not been documented in prior research.

The results show that the nine categories listed in Figure 1 are more sophisticated and multi-faceted than portrayed in previous research literature and require further evolution of conceptual and instructional models focused on the critical and interpretive demands of quantitative information (Gal, 2002; Weiland, 2019). The illustrations provided above, although referring only to two of the nine categories, demonstrate additional layers of knowledge and interpretive demands upon citizens, going well beyond those that are based on the analysis of single datasets or that focus on simple modelling as is the norm in regular instruction in statistics.

Further, media items about the COVID-19 pandemic present more *holistic demands*, including that citizens:

- *Combine or integrate* multiple ideas that are normally taught separately in mathematics and statistics,
- Grapple with *many literacy demands* (e.g., comprehension of written and spoken texts by journalists and experts)
- Be at ease with notions of *vagueness and risk* associated with statistical and mathematical products in the media.

Taken together, our findings indicate the need to take a fresh view of the knowledge and capabilities regarding statistical and mathematical knowledge that citizens should possess in order to understand essential information communicated in StaMPs within the media, which have direct bearing on their health and well-being. What shapes the public's perceptions of strength of evidence or scientific facts are critical issues that must be addressed in statistics education, as they impact on citizens' behaviors, such as compliance with official regulations.

It is important to note that numerous media items referred to terms and issues common in epidemiology or in official statistics (Radermacher 2020), such as the use of *statistical indicators* (which integrate information from multiple official sources into a synthetic estimate such as the R-rate); reporting of death rates as a proportion in relation to population size (e.g., deaths per 100,000 persons); use of 7-day averages to smooth unstable data; or comparison of death statistics in different years as a way to evaluate the excess deaths which may reflect the 'true' impact of the pandemic. Yet, such key indicators, that can be seen as part of *official statistics literacy* (Gal & Ograjenšek, 2017), are seldom addressed when learning mathematics and statistics.

The findings have many implications for theory and future research, as they raise new questions about what it means to be statistically or mathematically literate, and the need to study to what extend citizens actually possess the desired capabilities. The findings point to a need to expand existing models of statistical literacy and adult numeracy, and to further develop instructional efforts aimed at promoting the statistical literacy of students and adults-at-large. Further, based on the findings we argue for the need to accentuate efforts to link statistics education with conceptual models related to cross-curricular and global competencies, for example, 21st-century skills, and with broad views of adult numeracy which emphasize the need for criticality (Gal et al., 2020; Geiger et al., 2020). Indeed, the findings confirm the framework proposed by ProCivicStat Partners (2018) and elaborated by Gal et al (2021) to describe skills needed for engaging with *civic statistics* and their meaning of society and policy.

Many research and curricular questions emerge from the findings, such as: what extensions to current conceptualizations of the capabilities that all citizens should possess in order to be able to engage the wide range of STaMPs that can be found in digital and print-based media items; how to go beyond instruction in probability to understanding of vagueness of data or strength of evidence; or

how to develop students' ability to be critical regarding an expanded range of statistical and mathematical topics that are found in diverse real-world media items, encompassing both print-based and digital sources.

REFERENCES

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

Gal, I., & Ograjenšek, I. (2017). Official statistics and statistics education: Bridging the gap. *Journal of Official Statistics, 33*(1), 79-100.

Gal, I., Grotlüschen, A., Tout, D., & Kaiser, G. (2020). Numeracy, adult education, and vulnerable adults: A critical view of a neglected field. *ZDM Mathematics Education, 52*(3), 377-394.

Gal, I., Nicholson, J., & Ridgway, J. (Forthcoming 2021). A conceptual framework for Civic Statistics and its educational applications. In J. Ridgway (Ed.) *Statistics for empowerment and social engagement – teaching Civic Statistics to develop informed citizens*. Springer.

Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: A survey of the state of the field. *ZDM mathematics education, 47*(4), 531-548.

Geiger, V., Yasukawa, K., Fielding-Wells, J., Bennison, A., & Sawatzki, C. (2020). Facets of numeracy: Teaching, learning and practices. In J. Way, J. Anderson, J. Bobis, H. McMaster, K. Cartwright, & C. Attard (Eds.), *Research in mathematics education in Australasia: 2016-2019* (pp. 59-90). New York: Springer.

Hayne, J. (2020). Coronavirus 'nowcasting' modelling shows Australian case numbers continue to fall. *ABC*. Retrieved 24 April 2020 from https://www.abc.net.au/news/2020-04-24/government-releases-more-coronavirus-modelling-nowcast/12181500

Jablonka, E. (2003). Mathematical literacy. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick & F. K. S. Leung (Eds.), *2nd International handbook of mathematics education* (pp. 75-102). Springer.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks: Sage Publications.

Maass, K., Geiger, V., Ariza, M. R., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM - Mathematics Education, 51*(7), 869-884.

ProCivicStat Partners. (2018). *Engaging civic statistics: A call for action and recommendations*. (A product of the ProCivicStat Project). http://iase-web.org/islp/pcs

Radermacher, W. J. (2020). *Official statistics 4.0*: *Verified facts for people in the 21st century.* Cham, Switzerland: Springer.

Rahim, Z. (2020). Scotland recommends face coverings as cracks emerge in UK-wide approach to coronavirus. *CNN*. Retrieved 28 April 2020 from: https://edition.cnn.com/2020/04/28/uk/scotland-recommends-wearing-face-masks-gbr-intl/index.html

Steen, L. A. (1999). Numeracy: The new literacy for a data-drenched society. *Educational Leadership, 57*, 8-13.

Weiland, T. (2019). Critical mathematics education and statistics education: Possibilities for transforming the school mathematics curriculum. In G. Burrill and D. Ben-Zvi (Eds.), *Topics and Trends in Current Statistics Education Research* (ICME-13 Monographs, pp. 391-411). Springer.