Office for Statistics Regulation

CAMBRIDGE GRAND CHALLENGES

Internship Project

Statistical Literacy

Jessica McMaster

Cambridge Grand Challenges

September 2022

Contents

Executive summary

Introduction

The ability to understand and evaluate statistical information is crucial to navigate everyday life and the consequences of poor statistical comprehension are wide-ranging.

This report presents the findings from a literature review commissioned by the Office for Statistics Regulation (OSR) to establish the current landscape of statistical literacy research. This review provides a broad overview of the literature around statistical literacy and is not designed to be systematic or exhaustive.

Findings

Defining Statistical Literacy

The exact definition of the term statistical literacy is unclear, which may cause difficulties in both establishing the current levels within the general public and implementing strategies to improve it.

We found no consensus across definitions of statistical literacy, though there were some commonly used components, including foundational abilities, knowledge of statistical concepts, and critical thinking.

Differences between definitions are driven by contextual factors, and a definition broad enough to apply to all contexts may lose utility. An alternative approach for defining statistical literacy may be to consider first the context where the definition will be applied and then to specify the components required for statistical literacy in that context.

Knowledge

The majority of the evidence we identified related to the knowledge components linked to statistical literacy (e.g. literacy and numeracy) rather than statistical literacy directly.

We observed great variability amongst the general public in the skills linked to statistical literacy. Studies often found that skill level was influenced substantially by demographic factors such as age, gender, and education. More recent evidence is needed and we found no evidence that aimed to capture the multidimensional nature of statistical literacy.

Communication

We identified multiple recommendations on how to best communicate statistics to non-specialist audiences. The evidence for these recommendations varied from smaller-scale research studies to large-scale surveys, and recommendations from statistical bodies. The evidence base should be considered when deciding whether to apply these recommendations to statistical communication, with more weight given to recommendations when the study sample matches the target audience of the communication and when a finding has been replicated across multiple studies.

The examples of relevant OSR correspondence were largely aligned with the findings of the review although relevant correspondence could not be found for all topics.

Action

We identified a variety of initiatives in the statistical literacy space with different locations, target audiences and overarching aims. Many of the programmes may no longer be active and it may be beneficial to gain further information about these in particular. None of the work identified during the review aimed to connect these different initiatives. Therefore, sharing learnings across these groups may be a fruitful future endeavour.

Conclusion

This research was commissioned to shape OSR's future work on statistical literacy. The findings from the review provide an evidence base that will support OSR in developing its public position on statistical literacy and guide future regulatory work.

1 Introduction

1.1 Background

As society becomes increasingly data-driven, the ability to understand and evaluate statistical information becomes more crucial to navigating everyday life. Statistics are present in every facet of our lives including in important topics such as healthcare, economics, politics, and education. The inability to engage with this information can be severely detrimental to our success as a citizen within society (Ipsos, 2013).

Multiple organisations have stated the importance of ‘statistical literacy’. Valerie Isham, when President of the Royal Statistical Society (RSS), said that “Statistical literacy is an essential life skill: the need to make decisions based on numerical data confronts us all in every aspect of our professional and personal lives” (2012). Achieving and/or advocating for improved Statistical Literacy is included in the targets of the Government, in their 2021 Life Sciences Vision, and the UK Statistics Authority’s (UKSA) 2020-2025 strategy “Statistics for the public good”.

The consequences of poor statistical comprehension are wide-ranging. This can include having difficulty handling personal finances as well as being more susceptible to misperceptions and misinformation. Ipsos has been conducting a global research project since 2012 titled “Perils of perception” which records the gap between public perception and reality on a wide range of subjects. The overall findings of this project demonstrate widespread misperception on numerous topics including climate change, causes of death, and health. Bobby Duffy, while Managing Director of Ipsos Social Research Institute London, stated that one reason underlying these misperceptions is “our struggle with maths and proportions” (Duffy, 2016).

But what is statistical literacy? A 2017 review concluded that “statistics educators, statisticians and researchers around the world have not reached a consensus…and hence numerous definitions of statistical literacy abound” (Sharma, 2017).

Without a clear definition of statistical literacy, there are difficulties in both establishing the current levels within the general public and implementing strategies to improve it. When establishing strategies to improve statistical literacy, education-based initiatives may be considered. It is possible that this could place undue responsibility on the data consumer, we also do not know whether it is suitable or possible for everyone to invest the necessary time to develop complex statistical literacy. Evidence continues to emerge regarding how data *producers* can tailor their communication to achieve more widespread understanding and engagement for a population with varying levels of statistical literacy.

1.2 Aims of the review

This report presents the findings from a literature review to establish the current landscape of statistical literacy research. Research conducted in the area of statistical literacy is multifaceted. This report includes evidence on several specific topics.

The review was commissioned by the Office for Statistics Regulation (OSR), the regulatory arm of the UK Statistics Authority, to help shape its future work in this area.

The first section of this report explores how a consensus has not been reached on a single definition of statistical literacy. The section includes a review of the definitions that have been proposed thus far, to determine commonalities and differences, as well as the reasons underlying differences between definitions. The first section also covers the tools that have been applied in the measurement of statistical literacy as these require the developer to include an explicit or implicit definition of statistical literacy to determine how the tool should be constructed.

The second section outlines how researchers have attempted to measure components related to statistical literacy. This includes not only skills directly related to statistics, but also more foundational abilities such as general literacy and numeracy.

The third section provides evidence on the topic of statistical communication, including research on how to communicate effectively to non-specialist audiences to maximise understanding and engagement. There is a vast amount of research published on communicating different types of statistics, and this review is not intended to be exhaustive. Instead, the aim is to achieve a broad overview of research findings on several key topics.

Lastly, the report summarises actions taken in the UK and internationally to improve statistical literacy and statistical communication.

2 Defining Statistical Literacy

2.1 Definitions

Many definitions of statistical literacy have been proposed but there appears to be no consensus on how exactly statistical literacy should be defined (Sharma, 2017). The definitions identified through this review could be broadly grouped into either:

- Statistical literacy for students within an educational context **or**
- Statistical literacy for adults as data consumers within society

Context is just one of the factors driving the differences in definitions of statistical literacy. This section explores the factors causing variability in how statistical literacy is defined, and which components of statistical literacy are common across definitions.

2.1.1 Foundational Knowledge

At the most basic level, models of statistical literacy can include the foundational abilities required to engage with statistics (Chick & Pierce, 2013; Gal, 2002; Lehohla, 2002; ProCivicStat Partners, 2018; Schield, 1999). For example, Gal (2002) proposed that statistical literacy comprises both knowledge and dispositional elements (namely, critical stance, beliefs and attitudes). As well as knowledge of statistics, knowledge elements also include general literacy, understanding of context, critical questions, and numeracy skills. A number of definitions also mention the ability to “read” numbers and data which necessitates numeracy and literacy skills (Chick & Pierce, 2013; Lehohla, 2002; Schield, 1999). Although these foundational abilities are crucial in achieving statistical literacy, they are not included in all definitions. These low-level abilities are likely assumed to be implicit within the inclusion of more advanced skill sets.

The inclusion of foundational abilities in definitions of statistical literacy may be driven by contextual factors. The definition produced by Gal (2002) considers statistical literacy among adults who are data consumers in diverse life contexts. Whereas definitions produced in educational contexts (e.g. Ben-Zvi & Garfield, 2004), to develop assessments and teaching plans, may be narrower and foundational abilities such as general literacy and mathematical ability may be out of scope.

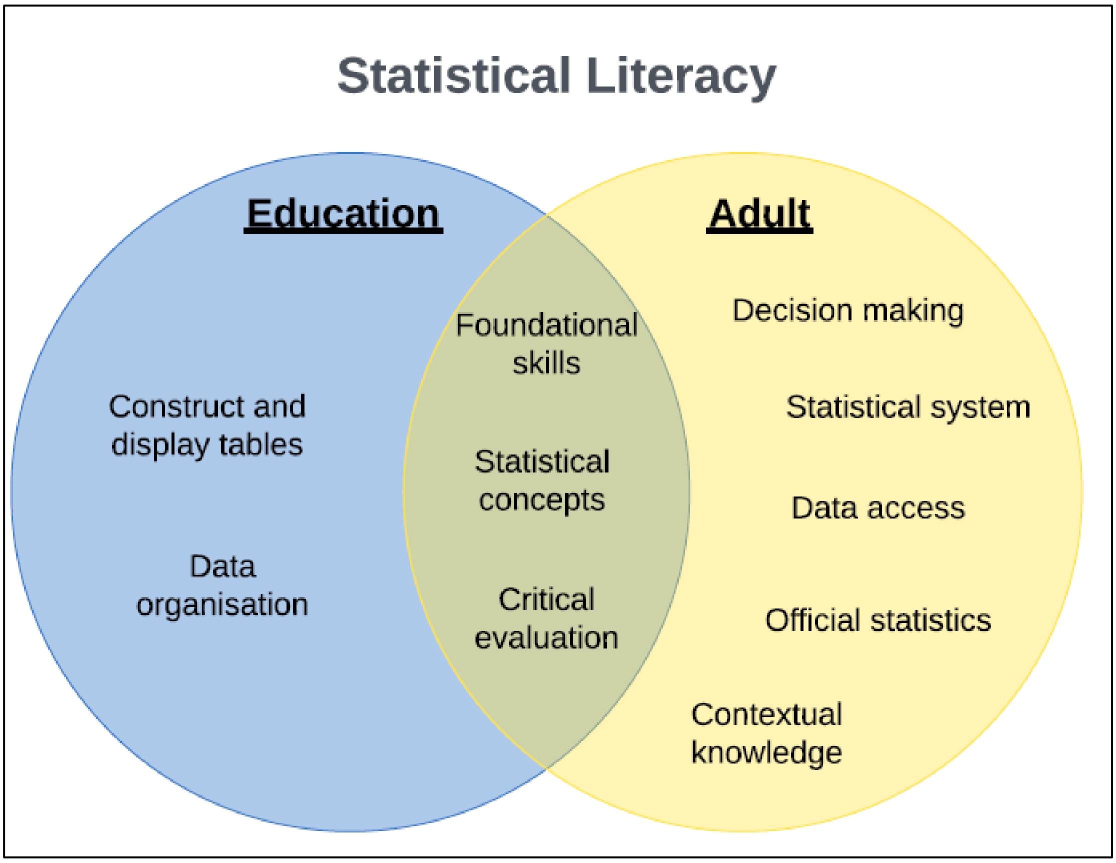

Figure 1. Statistical Literacy components grouped by context

2.1.2 Statistical Concepts

Knowledge and understanding of statistical concepts are ubiquitous in definitions of statistical literacy. As well as being one of the aforementioned knowledge elements in Gal's model of statistical literacy (2002), Ben-Zvi and Garfield (2004), make the distinction between statistical literacy, statistical reasoning, and statistical thinking. In this instance, statistical literacy includes an understanding of data organisation, table construction, probability as a measure of uncertainty, symbols, and vocabulary. Ben-Zvi and Garfield (2004) are statistical educators and therefore consider individuals as both data consumers and data producers. This is reflected in the knowledge they consider to be necessary for statistical literacy, namely practical abilities such as table construction.

When proposing the knowledge base required to engage with civic statistics, as part of the ProCivicStat project (a partnership between six different academic institutions to promote civic engagement), "statistics and risk", "models, patterns and representations" and "methodology and enquiry processes" were included (ProCivicStat Partners, 2018). Differences in the knowledge components specified as necessary to achieve statistical literacy may be prompted by two factors. Firstly, the ProCivicStat project considers individuals primarily as citizens and data consumers, without the data-producing role which would typically be considered within statistics education.

A second contributing factor relates to the types of statistics considered. Civic statistics, the focus of the ProCivicStat project, are “statistics about important societal trends and about topics that matter to the social and economic well-being of citizens” (ProCivicStat Partners, 2018). They assume that “citizens need to be aware of and critically understand statistics regarding past trends, present situations, and possible future changes in key social and economic areas such as demographics, employment, wages, migration, health, crime, poverty, access to services, education, human rights, and public expenditures” (ProCivicStat Partners, 2018). Civic statistics are described as distinctive in that they are typically: multivariate, aggregate, dynamic, communicated through rich text and data visualizations, and embedded in a social context. Overall, although statistical knowledge is a common element of statistical literacy models, the specific concepts that are included can vary depending on the setting and the type of statistics considered.

2.1.3 Critical Thinking

As well as being able to understand the statistical information presented, a consumer of statistical information is also required to take into consideration the quality of the information so that it can be appropriately incorporated into their thinking and decision-making. Within Gal’s statistical literacy model (2002), critical evaluation is captured in both the knowledge- and disposition-related elements. Knowledge of a set of critical “worry questions” is necessary to apply to statistical information to discern its quality (e.g., Is a given graph drawn appropriately or does it distort a trend in the data?). Furthermore, a critical stance is a dispositional element reflecting a person’s questioning attitude, including their willingness to apply the aforementioned worry questions to data that may be misleading or biased. Critical evaluation was also captured in other definitions of statistical literacy (Schield, 1999; Wallman, 1993). For example, Katherine K. Wallman, the President of the American Statistical Association, in 1993 produced a definition, based on her reading across many sources, which many recent models have been based upon: “Statistical literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives – coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions” (Wallman, 1993).

2.1.4 Additional Components

Beyond the components of statistical literacy identified as common across definitions, some elements are less frequently mentioned but are nevertheless necessary to engage with statistics in specific contexts. Additional elements which have been considered within definitions of statistical literacy include: incorporating statistical thinking into decision making (Wallman, 1993), ICT and search (ProCivicStat Partners, 2018), being able to access statistical reports and knowing where to source them, as well as wider knowledge of the statistical system which includes knowing where data comes from and how often it is released. The latter components are included particularly in the context of official statistics (Gal & Ograjensek, 2017).

2.1.5 Structure

As well as affecting the content of the definition, context may also impact the structure of the definition. In a review of definitions and models of statistical literacy (Sharma, 2017), Sharma highlighted that the hierarchical structure of the statistical literacy model provided by two statistical educators, Watson and Callingham (2003), may be due to its consideration of students who will undergo assessment. The hierarchy was established by Watson and Callingham for use in a student population, which at the lowest level (idiosyncratic) is associated with basic mathematical skills and advancement to higher levels is associated with more critical engagement with statistics in different contexts, understanding of uncertainty, and more complex statistical concepts (Table 1). Watson and Callingham evaluated their notion of statistical literacy using large-scale data from an 80-item questionnaire on statistical understanding. Their analysis of the questionnaire responses provided support for their conception of statistical literacy as a unidimensional construct with six levels of understanding. In contrast, Gal (2002, 2004) considers adult statistical literacy as a one-level construct containing the necessary components to be statistically literate.

Overall, a vast number of definitions and models of statistical literacy have been proposed. Although there is no consensus across these definitions, some common components have emerged. These are foundational abilities such as literacy and numeracy, knowledge of statistical concepts, and critical engagement with statistical information. Further components were identified which may be relevant to consider in specific contexts such as in official statistics (Gal & Ograjensek, 2017).

Table 1. Levels of Statistical Literacy (Watson & Callingham, 2003)

Level	Examples of Level Requirements
6) Critical mathematical	Critical engagement with context, use of proportional reasoning, appreciation of uncertainty.
5) Critical	Critical engagement in familiar and unfamiliar contexts (not involving proportional reasoning), use of appropriate terminology.
4) Consistent non-critical	Non-critical engagement with context, use of terminology, statistical skills e.g., mean, probabilities, graphs.
3) Inconsistent	Selective engagement with context, qualitative use of statistical ideas, recognise conclusions without justification.

2) Informal	Colloquial or informal engagement with context, basic calculations.
1) Idiosyncratic	idiosyncratic engagement with context, tautological use of terminology, and basic mathematical skills

2.2 Measurement tools

The lack of consensus regarding the definition of statistical literacy and other related and overlapping concepts such as data literacy makes it difficult to measure consistently (Open Data Institute, 2022). A number of tools to measure statistical literacy were identified during the review process which may shed further light on how to define statistical literacy.

Though all of the tools identified were developed to measure statistical literacy, context was once again an important factor. The context in which the tools were developed, as well as their wider purpose, substantially affected how they were constructed. For example, the Partnership in Statistics for Development in the 21st Century (PARIS21) Secretariat established a task team to develop a global indicator of statistical literacy. The Statistical Literacy Indicator (SLI) was developed to assess literacy at the country level and is based on articles produced by the countries' national newspapers (Klein et al., 2016; PARIS21, 2021). For this measure, statistical literacy is determined using a text mining approach to detect whether the media uses words representing the higher levels of the aforementioned statistical literacy hierarchy proposed by Watson and Callingham (2003; Levels 4-6 in Table 1). Level one here is the "Consistent, non-critical" use of statistics, Level two is "Critical" engagement with statistics and Level three is "Critical mathematical" engagement with statistics. A score is then estimated for each country, which is the percentage of national newspaper articles that contain references to statistics and use key words related to levels one, two and three summed (score ranging from 0 to 300). This tool has some clear limitations. One that was identified by the producers of the tool, was that in many developing countries the press releases from statistical agencies were used verbatim within their articles rather than editing them for their audience (Klein et al., 2016). This would artificially inflate the countries' literacy score. Another limitation is that it only factors in one type of national media, not including television or radio (PARIS21, 2016).

The review also identified two statistical literacy measures for use in a student population. Although, numerous measures were identified that capture statistical knowledge, the following measures have been developed to assess statistical literacy specifically. Detail is provided on the two assessments below which are the reasoning and literacy assessment (REALI; Sabbag et al., 2018) and Basic Literacy in Statistics (BLIS; Ziegler, 2014).

REALI (Sabbag et al., 2018) was developed for use in educational settings and estimates statistical literacy at the individual level, assessing performance on a series of questions. This assessment distinguishes between statistical literacy and statistical reasoning. Statistical literacy in this framework refers to knowledge of statistics and statistical reasoning is the ability to make more high-level connections between statistical concepts. Being critical of the statistics is not captured explicitly within either of these constructs; however, it is argued that the items used do "contemplate students' ability to be critical of statistical information" (Sabbag et al., 2018).

The definition of statistical literacy used in the construction of BLIS (Ziegler & Garfield, 2018) was "the ability to read, understand and communicate statistical information." BLIS is a 37-item assessment testing knowledge in areas such as data production, confidence intervals and regression. BLIS is once again more focused on detecting statistical knowledge rather than an individual's critical assessment abilities. Furthermore, they are both aiming to encompass the learning aims of introductory statistics courses.

The pattern we see across the tools captured in this review is that assessments and definitions that arise from an educational context appear to focus more on statistical knowledge. When the aim is to ascertain levels of statistical literacy within the general public and society, critical engagement is a much more central factor. Only a small number of tools were identified from the review therefore the strength of the conclusions are limited. Evidence from the use of these tools to estimate statistical literacy in the general public is discussed in the "Knowledge" section of this report.

2.3 Summary

Overall, the review captured various definitions of statistical literacy. In line with a prior review (Sharma, 2017), there appears to be **no consensus** on how exactly statistical literacy should be defined. The development process underlying these definitions was also often unclear. There are some **commonly used components** including **foundational abilities**, **knowledge of statistical concepts**, and **critical thinking**. Differences between definitions are especially apparent when comparing those produced for students and those for adults in society. A definition broad enough to apply to all contexts may lose utility. An alternative approach for future attempts at defining statistical literacy may be to consider first the context where the definition will be applied and then specify the **components required for statistical literacy in that context.** An example of this approach is displayed in Figure 1 where components of statistical literacy are grouped by context (Education and Adult).

3 Knowledge

There is limited evidence regarding the current level of statistical knowledge in the adult general population. As identified in the previous section, statistical literacy is also defined by several foundational abilities for which there is more evidence. In particular, this includes literacy and numeracy skills which have been assessed within multiple large-scale surveys of adult skills. This section summarises evidence on how researchers have attempted to measure components related to statistical literacy. This covers foundational skills, knowledge of statistical concepts, contextual knowledge, as well as revisiting the statistical literacy measures from the previous section.

3.1 Foundational Skills

The Skills for Life Survey (SfL2011), commissioned in 2011 by the Department for Business, Innovation and Skills, evaluated the basic skill level of adults aged 16-65 years in England (Department for Business Innovation & Skills, 2012). The SfL2011 respondents were recruited to be representative of 16–65-year-olds in England based on the proportion within each age band, ethnic group, and gender as well as the relative number who were disabled, employed or out of the labour market.

The skills assessed in SfL2011 included literacy, numeracy, and ICT skills. Each of these skills were pinpointed in the previous section as being foundational to achieving statistical literacy. Numeracy performance was assessed in 5,823 individuals and a skill level was subsequently assigned ranging from "Entry level 1 or below" to "Level 2+". Of those surveyed, 49.1% were Entry level 3 or below for numeracy. Entry level 3 is the national school curriculum equivalent for attainment at ages 9-11. Performance was higher overall in literacy with just 14.9% of the 5,824 adults surveyed scoring at Entry level 3 or below. Three practical skill areas of ICT were also assessed in over 2,220 respondents, these were word processing, email, and spreadsheets. More respondents were scored at the lower levels in spreadsheets with 66% scored at entry level 3 or below, in comparison 59% were scored at entry level 3 or below for word processing and 40% for emailing.

Individual-level variation in skill level could partly be explained based on demographic factors. For example, older adults performed considerably poorer than younger adults in ICT. Furthermore, females performed worse overall in numeracy than males. Other influential factors included first language, socioeconomic status, health, and ethnicity. The survey analysis also linked assessment performance to everyday life activities. For example, results indicated that responders that rated their numeracy performance as weak reported avoiding checking their bills and bank statements. Furthermore, they found that full-time workers had better numeracy than part-time workers.

The sample size for this survey was large and was designed to be representative of the population in England. The literacy and numeracy assessments had also been carried out in 2003 allowing observation of skill changes across that period at the group level. Overall, the results indicated a widely varied level of adult skills within the general population with individual differences partly driven by demographic factors. A notable proportion of respondents scored below Entry level 3 in numeracy which is likely to impact their ability to engage with statistical information in their everyday lives.

Foundational abilities were also assessed in the Programme for International Assessment of Adult Competencies (PIAAC). This is an international survey carried out every ten years in over 40 countries and involves interviewing 5,000 adults, aged 16-65 years, in each of the participating countries (Department for Business, Innovation & Skills., 2013). The skills assessed are literacy, numeracy and problem solving within technology-rich environments. The survey has had two cycles of data collection so far, with the results of the second cycle due to be published in 2024. Results from the first cycle of the survey, published in 2012, demonstrated that a substantial percentage of surveyed adults in most countries demonstrated the lowest levels of numeracy (8–32%) and literacy (5-28%) skills. In many of the countries surveyed, a high proportion of participants had low experience with ICT. Only 5.1% of adults, on average across countries, reached the highest skill level for problem-solving in technology-rich environments.

The mean proficiency score of just England-based respondents in literacy was not significantly different from the average of the other participating countries, whereas the proficiency score for numeracy was significantly below average. Around 10% of respondents in England reported a lack of experience or a lack of basic skills with computers. Consistent with the results of SfL2011, there is great variability in skill levels within the general population amongst the foundational skills required to achieve statistical literacy. This highlights the importance of tailoring statistical communication to different skill levels to reach a wider audience. This is discussed further in later sections of the report.

More recent evidence on numeracy skills in the general population was collected by the Financial Conduct Authority in their 2020 Financial Lives Survey (Financial Conduct Authority, 2021). This nationally representative survey is conducted approximately every two years, with the results of the first wave published in 2017. The 2020 survey involved around 16,000 interviews and as part of investigations into factors related to consumer vulnerability, data on numeracy skills related to financial concepts were collected. The results indicated that 34% of participants had poor or low levels of numeracy in this area. This assessment was based on just three multiple choice questions (Figure 2) with low numeracy indicated by answering none or one question correctly. Yet this survey does provide up-to-date evidence that the numeracy skills required as a member of society may be lacking within the general population causing quite widespread consumer vulnerability. Action to improve communication directed to consumers with varying skill levels has recently been carried out by the Plain Numbers Project (2021) which is discussed further in the "Action" section of this report.

Numeracy related to financial concepts questions:

NUM1. Suppose you put £100 into a savings account with a guaranteed interest rate of 2% per year. There are no fees or tax to pay. You don't make any further payments into this account and you don't withdraw any money. How much would be in the account at the end of the first year, once the interest payment is made?

NUM2. And how much would be in the account at the end of five years (remembering that there are no fees or tax deductions)?
1. More than £110; 2. Exactly £110; 3. Less than £110; 4. It is impossible to tell from the information given; 5. Don't know

NUM3. If the inflation rate is 5% and the interest rate you get on your savings is 3%, will your savings have more, less or the same amount of buying power in a year's time?
1. More; 2. The same; 3. Less; 4. Don't know

Figure 2. Numeracy questions used in the 2020 Financial Lives Survey (Financial Conduct Authority, 2021)

Research on numeracy was also conducted by Ipsos MORI, in partnership with National Numeracy and the Policy Institute at King's College London to mark National Numeracy Day in 2019 (National Numeracy, 2019). This involved an online survey, completed by 2,007 adults aged 16-75 years in the UK. The survey included five multiple choice numeracy questions designed to be roughly equivalent to a GCSE maths paper. The data from this survey was weighted to reflect the national population profile. The results indicated that 20% of the population scored 4 or 5 of the questions correctly which was described as roughly equivalent to a GCSE pass grade. They state that these results are in line with the numeracy findings from the aforementioned SfL2011 (Department for Business Innovation & Skills, 2012). Those in older age groups also tended to score higher than younger people.

3.2 Statistical Knowledge

One survey which did incorporate items around statistical information is the OECD's Programme for International Student Assessment (PISA). This survey has had multiple cycles, the first in 2000. The 2012 cycle involved data from around 510,000 students aged 15-16 years, across 65 countries, and is the most recent cycle to include data on statistical knowledge in the United Kingdom. The results indicated that the United Kingdom performed around the average in mathematics compared with the other participating countries (OECD, 2012).

In particular, the assessment area "uncertainty and data" tests knowledge of statistics and probability. Specifically, this content category is described as capturing the ability to detect and summarise messages within data as well as understanding the impact of variability. Of the four mathematics-based content categories that were assessed, including "uncertainty and data", "space and shape", "quantity" and “change and relationships”, the mean performance of respondents in England was highest for “uncertainty and data”. It was concluded that pupils in England had a relatively strong performance on the assessment questions related to probability in statistics. In the wider United Kingdom, all countries perform higher on the “uncertainty and data” subscale than their overall mathematics score. Wales’ scores on all four mathematical subscales were lower than the other UK countries. Overall, although this survey tested pupils aged 15-16 years it could be argued that this is representative of a large proportion of adults’ statistical knowledge as many do not complete further mathematics-based education beyond this point; although, this review does not consider evidence around the retention of knowledge over time. OSR evaluated reports of PISA data against the Code of Practice for Statistics in 2021. The findings involved several recommendations predominantly around including more information about potential limitations and sources of bias in the data. It is expected that this will be addressed in future reports of PISA findings.

Public understanding of statistics was also captured in a 2020 report published by the Economic Statistics Centre of Excellence (ESCoE) titled “Public Understanding of Economics and Economic Statistics”. As mentioned in the title, this survey focused on economic statistics in particular. A series of research involving 12 focus groups and an online survey of UK adults explored understanding in areas such as inflation, unemployment, and gross domestic product (GDP) (Runge & Hudson, 2020). Overall, the results indicated that public understanding is greatest for areas where personal relevance is perceived to be higher such as inflation and interest rates, rather than GDP. Analysis of differences in results due to demographic factors suggested that older male participants of higher educational attainment and socioeconomic status were more knowledgeable, confident, and interested in economic statistics. Economists highlighted the importance of engaging with the public to gain insight that would improve their methods of communication. This was addressed in a follow-up study which is discussed in the next section of the report (Runge & Killick, 2021).

The report outlining this work also included a literature review around the subject area of public understanding of economic statistics. The review made the distinction between ‘top-down’ approaches to studying public understanding, whereby understanding of economic concepts is tested using the definitions applied by the economists, and ‘bottom-up’ approaches where participants are asked open-ended questions to gain insight into how the public describes their own understanding of economic concepts. The former “top-down” approach may not take into account understanding held by the public that does not fit into the set definitions used. One of the main findings of the bottom-up literature reviewed was that people often understood the economy using metaphors (e.g. as a machine). The use of metaphors may result in the public simplifying and misunderstanding complex topics. They may also be erroneously confident in their understanding of the topic by assuming that the economy will follow the behaviour of the metaphor they utilise (see: Runge & Hudson, 2020).

3.3 Contextual Knowledge

Understanding and interpreting statistics involves not only skill-based knowledge such as the components captured above but also wider contextual knowledge to put statistics into perspective. This type of knowledge was assessed in a widescale study of European's knowledge of official economic statistics (Vicente & López, 2017). Data was collected as part of the Eurobarometer 83.3, carried out on behalf of the European Commission, and including ~28,000 individuals aged over 15 years in 28 EU Member States.

Respondents were asked questions about economic statistics within their own country (e.g., "*Do you think that, in your country, the inflation rate in 2014 was higher, lower or equal to the rate in 2013?*"). Around a quarter of people in the United Kingdom responded "don't know" about each of the official figures which were growth rate, inflation rate and unemployment rate. This suggests a substantial proportion of the population may lack the necessary contextual knowledge to put economic statistics into perspective.

The results of this study highlight the importance of including important contextual information in statistical communication to engage with a wide-ranging audience of different knowledge levels.

3.4 Statistical Literacy

This section of the report has so far captured evidence on how researchers have attempted to establish knowledge, within the general public, on topics which were identified in the first section of the report to be linked to statistical literacy. As previously mentioned, researchers have also developed tools to measure statistical literacy as a whole. This section summarises instances of when these tools have been applied to non-specialist audiences.

The Statistical Literacy Indicator, introduced in the previous section, was applied to evaluate the use of statistics in national newspapers (PARIS21, 2021). The results for the United Kingdom indicated that 18.8% of newspaper articles engaged with statistics at Level 1 in 2020 (non-critical). This is higher than in previous years (starting from 2016) when the percentage was typically between 10-13%. Only 0.3% of articles had text classed as Level 2 (critical) engagement with statistics, which was marginally lower than in previous years where the percentage ranged from 1.1-1.5%. A similar pattern was observed for Level 3 (critical mathematical) where the percentage was very low at 0% in 2020 but had been somewhat higher in previous years (0.9-1.9%).

Overall, the results suggest that mention of statistics was higher in newspaper articles in 2020, the beginning of the COVID-19 pandemic. The use of critical language related to statistics did not rise from previous years and was at a very low level overall. Although, these results are described as preliminary before checking by analysts of National Statistical Offices (NSOs).

When establishing the REALI assessment, data were collected for 671 students at the graduate or undergraduate level (Sabbag et al., 2018). As previously mentioned, this assessment distinguishes between statistical literacy and statistical reasoning. The results indicated that the mean scores out of 20, for the statistical literacy and statistical reasoning subscores, were 13.15 and 11.01 respectively. The knowledge of students at college or university is unlikely to be representative of the general population. These data were collected to evaluate the questionnaire in students, who are the target for this tool and so meets that aim but is less applicable for the wider population.

This limitation is also relevant to the data from the BLIS assessment (Ziegler & Garfield, 2018), which was applied to 940 college students who scored on average 21.41 out of 36. These data were collected to evaluate the assessment for students and the results do not apply to the wider public. Both, the REALI and BLIS assessments have not, as far as was identified in the review, been applied outside of a student population. Based on the conclusions of the previous report section, it is likely that there would be different areas of interest when assessing statistical literacy in adults as these two education-based assessments appear to focus predominantly on knowledge of statistical concepts.

Statistical literacy was also assessed in an international survey conducted by the W. M. Keck Statistical Literacy Project in 2002 (Schield, 2006b; described further in the "Action" section). This statistical literacy survey focused on the skill of reading graphs and tables of rates and percentages in 191 participants made up of US college students, college teachers (worldwide) and data analysts in the US and South Africa. The results indicated high error rates, for example, when participants were asked a question comparing two percentages there was a high and similar error rate among students and college teachers (student: 82%; teacher: 81%) while data analysts performed better (60%). Although this is described as a statistical literacy survey it is focused on just the ability to read graphs and tables of rates and percentages. It is also applied in a relatively small sample and was conducted twenty years ago. Therefore, the findings of this study should be interpreted with caution and should not be generalised to the current population.

3.5 Other knowledge

The previous section of this report ("Definitions") highlighted that there are a wide range of concepts linked to statistical literacy, particularly when considering official statistics. This included being able to access statistical reports and wider knowledge of the statistical system.

Research was published this year, which was commissioned by the UK Statistics Authority, on the topic of public confidence in official statistics (Butt et al., 2022). Results are based on a survey of adults aged over 18 years, recruited to be representative of adults in England, Wales and Scotland, that was conducted by the National Centre for Social Research (NatCen). This survey has been conducted regularly since 2004 and the results of the 2021 survey showed that 75% of respondents (other than "don't know" or no response) had heard of the Office for

National Statistics (ONS) and 36% reported using ONS statistics. Furthermore, 64% of respondents thought official statistics were easy to find and 67% thought they were easy to understand. People were also asked how well they knew the UK Statistics Authority (the Authority) and the Office for Statistics Regulation (OSR). The results indicated that 48% of respondents had heard of the Authority (2% knew it well) and 41% had heard of OSR (2% knew it well). Overall, the survey indicates some awareness of statistical bodies in the UK. It also indicated that the majority of the respondents agreed that official statistics are easy to find and understand. A large minority disagreed with these statements. People with higher levels of education were more likely to agree and so were people aged 35-44. Whereas people aged over 65 were more likely to disagree.

3.6 Summary

In this section of the review, great **variability** was observed amongst the general public, in the skills that are linked to statistical literacy. Studies often found that skill level was influenced substantially by **demographic factors** such as age, gender, and education. **More up-to-date, large-scale surveys are now required** to provide ongoing insight into the knowledge level of the population. **No evidence was found that aimed to capture the multidimensional nature of statistical literacy.** This could be due to the breadth of statistical literacy which makes it difficult to define, as identified in the previous section of this report ("Definitions"). Overall, this section highlights the importance of generating statistical information that can be understood by audiences with varying skill levels. The factors to consider when developing statistical communication, to achieve greater understanding and engagement, are explored in the next section of this report.

4 Communication

This section summarises the evidence identified on the topic of statistical communication. In particular, on how to communicate effectively to non-specialist audiences to maximise understanding and engagement.

Due to the breadth of this evidence, it has been broken down into categories covering a wide range of factors, from considering the target audience, to how to communicate uncertainty within statistics.

4.1 Audience

The first factor to consider when developing statistical communication materials is the target audience (UNECE, 2009). In a review of risk and uncertainty communication published by David Spiegelhalter in 2017 (Spiegelhalter, 2017), it was recommended to consider the target groups of the communication and identify their needs, beliefs, and skills. This recommendation followed a recurring finding within the review that the best approach to communicating information can vary substantially depending on the characteristics of the audience.

A study comparing understanding of health-related statistical information, communicated using either numerical or graphical representations, found that the result was dependent on the graph literacy of the participant. In a study by Gaissmaier et al., (2012), each participant's graph literacy was assessed using a scale which assessed an individual's understanding of health-related information conveyed using graphical representations (Galesic & Garcia-Retamero, 2011). Participants with high graph literacy demonstrated greater comprehension and recall of the statistical information when presented with a graphical representation, whereas the reverse result was observed for adults with low graph literacy scores (Gaissmaier et al., 2012). Similarly, another study found that levels of numeracy had a significant impact on how participants perceived risk communicated either via frequencies or percentages (Peters et al., 2011). In particular, less-numerate participants presented with risk information in a percentage format perceived medication as less risky than when presented with the same information as a frequency. Whereas highly numerate participants perceived similar risks in both formats. These studies highlight the importance of considering the characteristics of the target audience when designing statistical communication materials.

Although some research has already been conducted on characteristics to consider, such as graph literacy (Gaissmaier et al., 2012), and numeracy (Kreuzmair et al., 2016; Peters et al., 2011), it is also recommended to carry out your own development process of testing and evaluating materials within target groups (Spiegelhalter, 2017). This process was undertaken in the development of infographics to communicate personal risk from COVID-19 (Freeman et al., 2021). This began with an initial round of qualitative interviews both with members of the general public and with primary care physicians, followed by further rounds of refining the material design in qualitative interviews as well as quantitative experiments in large samples of participants.

This process provided valuable insight such as finding out that the general public does not consider their risk from COVID-19 in a numerical way. A consequence of this finding was a recommendation to translate numerical risk to match more closely with the user's own subjective experience. This included a more qualitative description of risk including a reference point of a person with known health risk factors of a particular demographic. Subsequent research further highlighted that using illustrative personas with clear risk factors provided useful context to enable users to understand their level of risk from COVID-19 (Freeman et al., 2021).

Following a series of workshops involving economists and the general public, it was also recommended to conduct future public engagement studies with economists and the general public to discuss economic issues and gain information on how to improve communication with the public (Runge & Killick, 2021). This research focussed on economic statistics, but this recommendation may be generalisable to other areas of statistics.

As well as aligning communication with the audience’s thinking on the topic, it has also been recommended to use language suitable for the target audience (Spiegelhalter, 2017). Understanding may also be improved by emphasising the relevance of the information to the audience's lives. Public understanding of Economics and Economic Statistics was captured in a 2020 report published by the Economic Statistics Centre of Excellence (Runge & Hudson, 2020). A series of studies involving 12 focus groups and an online survey of UK adults explored understanding in areas such as inflation, unemployment, and GDP. Overall, the results indicated that public understanding was greatest for areas where personal relevance is perceived to be higher such as inflation and interest rates rather than GDP (ESCoE, 2020). One approach to enhancing the relevance of statistics could be breaking down high-level (e.g. country-level) statistics into local or demographic-based statistics so that users can access those most relevant to them. The tailoring of statistical information to enhance relatability was shown to be effective in improving public comprehension and trust in research conducted by the Bank of England and the Behavioural Insights Team. This study is discussed further in subsequent sections (Bholat et al., 2018).

4.2 Context

The second theme that emerged from this part of the review was the importance of providing contextual information with statistics. Broadly this involves ensuring statistics are always contextualised so that audiences can comprehend their significance (BBC Trust, 2016). This also encapsulates more specific recommendations such as including baseline risk when discussing changes to risk (BBC Trust, 2016), emphasising personal relevance to the audience (Bholat et al., 2018; Runge & Hudson, 2020), and establishing a narrative within the communication materials to improve understanding (Runge & Hudson, 2020; Spiegelhalter, 2017; UNECE, 2009). This narrative should be vivid but should not cause undue affective responses (Spiegelhalter, 2017).

Another important aspect of a statistic's context is its source. In a series of recommendations, developed by authors at the Norwegian Institute of Public Health on how to communicate evidence-based information about the effects of healthcare interventions, it was advised to provide relevant background information (Oxman et al., 2020). This could include how the information was put together, what it is based on, the people who put it together and whether they have any conflicts of interest. It was stated that this would allow users to understand why, and if, information is trustworthy (BBC Trust, 2016; Oxman et al., 2020; Runge & Hudson, 2020).

Lastly, the importance of choosing appropriate comparators was highlighted by several sources. This includes Spiegelhalter (2017), who highlighted that comparators are useful when communicating risk to people with low numeracy, but some comparators can be associated with an emotional response. For example, the risk of being struck by lightning is a poor comparator because it is newsworthy and is therefore perceived as more likely than it is. Therefore, its use as a comparator would lead to misperception.

OSR Correspondence on Context

OSR challenged the UK Health Security Agency's COVID-19 vaccine statistics for using an inappropriate comparator when comparing COVID-19 rates in vaccinated and unvaccinated populations. These populations have known differences, which may include differences in the likelihood of coming forward to be tested (1/11/2021).

Click here to see correspondence in full

4.3 Language

The next topic surrounds language use within statistical communication. Overall, the message observed consistently across several articles is to use simple and easy-to-understand language (Bholat et al., 2018; Oxman et al., 2020; Spiegelhalter, 2017; The Plain Numbers Project, 2021; UNECE, 2009). In particular, one should avoid using technical language or 'jargon'. The level of technical language that should be used should be determined by the intended target audience. Language use should also be consistent (Oxman et al., 2020) and only necessary information should be included (Spiegelhalter, 2017).

OSR Correspondence on Language

In a compliance check of statistics from the Scottish June Agricultural Census, the use of jargon was challenged with the following recommendation:

"*We encourage you to minimise the use of jargon and add definitions where possible to help a wide range of users understand the statistics*." (28/3/19)

Click here to see correspondence in full

4.4 Format

Another aspect of statistical communication that should be considered is the format of the statistical information itself. One area in which there is some debate is the use of frequencies versus probabilities, particularly when communicating risk.

In the context of diagnostic or screening tests, the evidence seems to be in favour of natural frequencies (Akl et al., 2011; Hoffrage et al., 2000). Here, a natural frequency refers to the joint frequency of two events (e.g., the number of people who have a disease and the number who would have a positive test result using a particular screening tool). Findings from a systematic literature review (Akl et al., 2011) suggested that natural frequencies are better understood than probabilities, this result was also argued in an earlier article published in 2000 (Hoffrage et al., 2000).

This result may be specific to the context of presenting joint probabilities. In an aforementioned study developing infographics to convey the personal risk of COVID-19 (Freeman et al., 2021), users showed lower variation in their estimation of risk when shown percentages. Probabilities also had the consequence of lowering the user's perceived level of risk than when shown the same information as a frequency. Overall, this study concluded that percentages conveyed risk most clearly but resulted in an underestimation of risk. Therefore, the recommendation was to present both percentages and frequencies for balance.

Spiegelhalter (2017) argues that being clear is the most important consideration above the choice of format. If using a frequency format, then he argues that there are choices to be made about the denominator due to ratio bias (a larger numerator suggests a larger risk). The author further recommends choosing a frequency format with a clear reference class, using “1 in X” can be seen as suggesting higher risk when expressed with a higher numerator particularly when the user has a low educational level.

Additional, format-based recommendations include reporting absolute effects (Oxman et al., 2020), and being cautious of using percentages when the numbers are small to avoid misinterpretation (Home Office, 2013).

4.5 Framing

Minor changes in the wording of information can have important implications on how it is perceived. Framing is an important issue in healthcare and one clear example of this is the choice of “survival rate” versus “mortality rate” when conveying risk information. One study using this language to investigate the impacts of positive and negative framing, found that negative framing (“mortality rate”) was associated with greater perceived uncertainty in the risk information as well as greater perceived risk (Freeman et al., 2021). Positive framing (“survival rate”) was, perhaps unsurprisingly, relatively well-liked by users and perceived as less concerning than negative framing, although it was, in some cases, associated with poorer comprehension of risk. Though this is described as an exploratory finding and should therefore be treated with caution before this result has been fully replicated (Freeman et al., 2021). Another study similarly found that positive framing of risk information ("90% of patients **do not** get a bad blistering rash") resulted in a lower perceived risk than negatively framed information ("10% of patients get a bad blistering rash"; Peters et al., 2011).

It was recommended in all of the studies that discussed framing within this review (though with some common authors), that both positively and negatively framed text should be presented to avoid unintended biases (Freeman et al., 2021; Peters et al., 2011; Spiegelhalter, 2017). It has also been recommended to use visualisations of part-to-whole comparisons of both positive and negative outcomes (Spiegelhalter, 2017).

4.6 Trust

Trust is another consideration which emerged from the review, specifically the recommendation to include information in statistical communication that helps users to understand whether it should be trusted (Oxman et al., 2020). This has clear links to the prior topic of context as knowing how the information was put together is an important factor when judging its reliability. As well as contextual information, detail should also be provided on any limitations or quality issues affecting the data (Home Office, 2013).

Relevant information may include who produced the statistical information and how it was produced (Oxman et al., 2020; UNECE, 2014). The potential implications of a lack of clarity on the source of statistics were highlighted in the previously mentioned report on public understanding of Economic Statistics (Runge & Hudson, 2020). The focus group research conducted in the development of that report revealed that the participants often erroneously associated economic data with the government and the politicians they perceived presenting them in the news. This association resulted in some of the focus group participants communicating a lack of confidence in the accuracy and reliability of the economic figures presented. Economic statistics, such as unemployment and inflation rates are produced by the Office for National Statistics (ONS) the UK's national statistical institute which is independent of government. If this was expressed more clearly perhaps confidence in these statistics would be raised, highlighting the importance of this information. In a series of workshops involving economists and members of the general public, economists also argued that distrust, due to suspicion that the government manipulates unemployment statistics, could be somewhat addressed by highlighting the ONS' independence of government (Runge & Killick, 2021).

The focus group involved in ESCoE's 2020 report expressed that distrust in economic statistics was also caused by the view that economic topics were discussed in an inaccessible way, using economic jargon (Runge & Hudson, 2020). This further highlights the importance of using language suitable for the target audience which was first considered in the "Language" section of this report.

Research conducted jointly by the Bank of England and the Behavioural Insights Team (BIT), found that using more relatable language in their communication of information related to inflation led to greater understanding (Bholat et al., 2018). Relatable language included using more first and second-person pronouns and using more "day-to-day" terms rather than technical language. Using relatable language also led to small but statistically significant increases in the participant's trust in the information and perception of the bank.

Another connection between language choice and trust is deciding whether to use numbers or words. In a review of risk communication, Visschers (2009) included the recommendation that both numerical and verbal probability information should be included when communicating risks. This is because people prefer the accuracy of numerical information, trusting and understanding it more, but pass it on using evaluative words. Therefore, people require both formats to be fully equipped with the information they will need to engage with the statistics (Visschers et al., 2009).

As well as providing information to allow users to judge whether evidence is trustworthy, the quality of evidence can also be stated explicitly. One study explored the effects of communicating different levels of evidence quality on the perceived trustworthiness of the evidence (Schneider et al., 2021). Results indicated that when participants were told that evidence was of low quality it was perceived as less trustworthy. When users were told the evidence was of high quality, the perceived trustworthiness of the information was not significantly different than when no information at all was provided on evidence quality. This may indicate that when no information is provided regarding evidence quality users assume that the evidence is of high quality. This has clear potential negative implications of providing no information about evidence quality, particularly when presenting evidence of low quality.

Lastly, evidence suggests that trust is detrimentally impacted when an outcome is different from the expectation (Runge & Killick, 2021; Vicol, 2020). This highlights further, the importance of providing sufficient information so that the reasons underlying an unexpected result are understood.

OSR Correspondence on Trustworthiness

OSR challenged the use of unpublished data in the Government COVID-19 press briefing. An estimate of the cost of the UK Health Security Agency's Test and Trace programme for January 2022 was included without an appropriate explanation of context and sources (4/3/22).

Click here to see correspondence in full

4.7 Uncertainty

Communication of uncertainty is a major topic in the literature and admitting uncertainty in statistics is one of the recommendations made by David Spiegelhalter in his 2017 review of risk communication (Spiegelhalter, 2017). Oxman et al., (2020) also recommended that the certainty of evidence should be explicitly assessed and reported.

An ESCoE survey regarding the communication of uncertainty, particularly around the U.K GDP, found that communication of uncertainty information was useful in ensuring that the public does not take GDP estimates at point value while not decreasing trust in the data. In particular, it is recommended to communicate uncertainty quantitatively (e.g. using intervals, density strips and bell curves; Galvão et al., 2019).

More recent ESCoE research, based on 20 qualitative interviews with members of the UK public, found that perceptions of uncertainty can vary dependent on the statistic being communicated (Runge, 2021). In this study, uncertainty was presented around GDP and unemployment statistics, and these were discussed in semi-structured interviews. Participants engaged more with uncertainty information about unemployment statistics. Whereas some respondents stated that, for GDP, they would have preferred just the statistic without the uncertainty information. This is aligned with recommendations to tailor messaging for audiences with different information needs and interest levels. The interviews also covered how participants felt about the underlying causes of uncertainty. For unemployment statistics, information around the method of collecting these data challenged common assumptions. In this instance, a greater amount of method-based information may be useful to aid understanding.

Further evidence on how uncertainty should be communicated was summarised by FullFact in a 2020 briefing "How to communicate uncertainty". They collated several recommendations including being transparent, being specific about what exactly is uncertain, indicating uncertainty in existing data using numerical ranges in brackets after the main value, and when making future predictions using verbal expressions supplemented with numerical probability ranges citing underlying data where possible (Vicol, 2020). Furthermore, they identified research indicating that verbal indicators of uncertainty are more open to interpretation (see also: Dhami, 2018) and can be interpreted cumulatively. For example, if an outcome is described as "likely" across multiple sources it may be perceived by some as "very likely" (Mislavsky & Gaertig, 2022; Vicol, 2020).

A comparison of verbal and numerical uncertainty indicators was also included in a recent study investigating outcomes such as cognition and trust (Van Der Bles et al., 2020). The results indicated that the effects of communicating uncertainty were highly dependent on the format of the uncertainty indicator. Communicating uncertainty with verbal quantifiers (e.g., "it could be somewhat higher or lower") led to small significant declines in the perceived reliability of the message and the trustworthiness of the source of the information. Conveying uncertainty numerically was not associated with these outcomes (Van Der Bles et al., 2020).

Overall, it has been recommended to use both verbal and numerical indicators alongside each other and there is even advice on how they should be ordered.

Guidance on Communication of Uncertainty in Scientific Assessments from the European Food Authority recommends that when presenting both verbal and numerical indicators of an approximate probability, numerical indicators should be presented first as in English this order is interpreted more consistently (EFSA, 2019). They also recommended providing information on sources of uncertainty estimates for an informed audience.

Beyond the question of numerical or verbal uncertainty indicators, how should uncertainty be presented? One study recommended using a probability function to communicate uncertainty, where relevant, to non-experts (Greis et al., 2015). This finding was based on experimental data with a relatively small sample, investigating behaviour based on expected rainfall statistics with uncertainty communicated in a variety of ways.

For additional information regarding the communication of uncertainty, the OSR has recently undertaken a review into "Approaches to communicating uncertainty in the statistical system". This review captures the guidance available to statistical producers, past recommendations made by the OSR, and examples of good practice when presenting uncertainty.

OSR Correspondence on Uncertainty

In the Rapid Review of a weekly Public Health Scotland COVID-19 and Winter Statistical Report, the lack of uncertainty information was challenged by OSR and the potential resulting misuse and overinterpretation of the data was highlighted. Providing prominent explanations of the uncertainty and caveats around the statistics was suggested as one solution (11/2/22)

Click here to see correspondence in full

4.8 Visualisations

Visualisations have also been discussed more broadly, including recommendations from Spiegelhalter (2017) in the aforementioned review of risk communication. This includes the recommendation to avoid "chart junk" such as three-dimensional bar charts (Spiegelhalter, 2017; see also: UNECE, 2009). This recommendation was made when considering uncertainty communication but appears to apply to other statistical communication. Similarly, Spiegelhalter advises caution when incorporating interactivity or animation into statistical communication as it may introduce unnecessary complexity rather than being beneficial to comprehension of the statistical message.

Further recommendations regarding visualisations, from the same publication, include using multiple formats to suit different audience groups, adding numbers and words to graphs to aid comprehension, as well as useful and clear narrative labels.

Overall, the use of visualisations was shown to be potentially beneficial in a study involving the Bank of England (Bholat et al., 2018). Their visual summary of statistical information related to inflation, including engaging icons and charts, was related to higher comprehension compared to their traditional text-based communication. Their visual summary also used simpler language so the sole effect of using visualisations was unclear.

OSR Correspondence on Visualisations

OSR frequently provides positive feedback when visualisations are applied to aid understanding of statistics. One example is provided below:

Following a compliance check of the Ministry of Defence's (MOD) Armed Forces Continuous Attitude Survey statistics against the Code of Practice for Statistics, OSR commended the movement from using individual tables to visualisations to portray the data. OSR noted that this helps make technical data more engaging for non-technical users

Click here to see correspondence in full

4.9 Summary

Overall, this section of the review identified a **wide range of recommendations** on how to best communicate statistics to non-specialist audiences. The basis of these recommendations varied from smaller-scale research studies to large-scale surveys, and recommendations from statistical bodies. When deciding whether to apply these recommendations to statistical communication the **evidence base should be considered**, with more weight given to recommendations when the study sample matches the target audience of the communication and when a finding has been replicated across multiple studies.

This section also highlighted examples of when OSR have published correspondence in the areas being discussed. Where examples were identified, these appeared to be **aligned with the findings of the review**. Examples were more available on broader recommendations (e.g., uncertainty should be reported) and were not always found for more specific recommendations (e.g., reporting both verbal and numerical uncertainty indicators). This may be because these more specific recommendations have less of an evidence base.

5 Action

In Ferligoj's (2015) paper titled "How to improve statistical literacy", the actors that can contribute to better statistical literacy include educational institutions, statistical offices, statistical associations, and the media. Furthermore, recommendations were made on action to be taken which included: advising different segments of the population separately about the proper use and interpretation of statistical data as well as presenting typical abuses or misunderstandings of statistical concepts and data.

This section of the review encompasses action that has been taken, in the UK and internationally to improve statistical literacy and statistical communication.

5.1 United Kingdom

The Royal Statistical Society (RSS) launched the *getstats* campaign in 2010 which had a 10-year goal of improving statistical literacy in the UK. It aimed to do this by providing training and running workshops to statistics communicators, such as journalists, MPs, and their research staff. There would be wider benefits from RSS publishing the outcomes and learnings from this campaign.

In 2021, a user engagement strategy for "ensuring official statistics meet society's needs" was announced by the Government Statistical Service for the producers of official statistics within the UK. The three main goals of this strategy are: building collaboration across public sector bodies, developing the capabilities of statistics producers to understand the audience, and strengthening user engagement culture. Furthermore, the phases of the strategy include showcasing and building on existing good work as well as establishing baseline activity, establishing the USER hub to lead the way on user engagement support packages, and lastly, depending on statistics producers to use the USER hub to improve engagement. This strategy is still in its early stages and is expected to be implemented over four years (2021-2025).

The Winton Centre for Risk and Evidence Communication, the source of a number of the studies mentioned in this report, have established multiple online tools available on their website (https://wintoncentre.maths.cam.ac.uk/) that incorporate some of the learnings of this research. These include:

- **Predict websites:** Tools to allow clinicians to compare different treatment options for breast cancer (https://breast.predict.nhs.uk/) and prostate cancer (https://prostate.predict.nhs.uk/) both individually, and with patients during their appointments. They have incorporated uncertainty in their estimates using numerical ranges, which is in line with their previous research highlighting that this approach is not harmful to perceived trust in the data (Van Der Bles et al., 2020). Furthermore, work is ongoing to investigate the effect of presenting the statistics in different formats on people's understanding of the data.

- **RealRisk** (realrisk.wintoncentre.uk): An online tool to assist press officers and journalists in communicating risk statistics in health and social areas. Specifically, it can be used to convert risk originally conveyed as relative risk, odds ratios, or hazard ratios into absolute risks. As mentioned in the "format" section of this report, absolute risks are often better understood by the general public (Oxman et al., 2020).

- A tool **communicating risk from different COVID-19 transmission routes**. Statistical information is presented in different ways for different users by integrating information boxes into their infographics that communicate COVID-19 transmission risk so that more interested/advanced users can find out more about the certainty of evidence. Colour bars were also used rather than numerical estimates of risk to limit the overinterpretation of precision in these statistics (Rutter et al., 2021).

The Plain Numbers Project (https://plainnumbers.org.uk/) was established in 2021 to support consumers who struggle with numeracy. Plain Numbers has worked with the Bank of England as well as firms such as ClearScore and Direct Line to make changes to their communication materials to improve consumer comprehension. The approach is based on three key principles, namely "numbers themselves" which refers to aligning the use of numbers with how humans think which is in stories. The second principle is "numbers in context" which recommends assuming little contextual knowledge from the consumer as well as removing jargon. The third principle is "how we think" and this is related to considering how the information to be presented should be processed by the consumer. In initial studies the Plain Numbers approach was successful and across firms, the approach doubled the number of people who understood customer information (The Plain Numbers Project, 2021).

5.2 International

As part of the work programme of the Conference of European Statisticians, a Steering Group on Statistical Dissemination and Communication, supported by the United Nations Economic Commission for Europe (UNECE), produced four guides to help those communicating statistics to non-statisticians. The first guide "A guide to writing stories about numbers" (2009) was prepared by a group of international experts. The audience of the guide are managers, statisticians, and media relations officers. The guide begins by emphasising the importance of providing a statistical story. The guide provides recommendations for writing with a focus on increasing engagement by being "newsy". Overall, the guidelines are quite consistent with others in this area. Most relevant is the fourth guide, titled "How to improve statistical literacy: A guide for statistical organizations" which covers statistical literacy not only in the general public but also in decision-makers and within statistical organisations (UNECE, 2014). This guide includes a summary of several international groups working in the area of statistical literacy.

This summary included the International Statistical Literacy Project (ISLP) (https://iase-web.org/islp/) which was initiated by the International Association for Statistical Education (IASE), the education section of the International Statistical Institute (ISI) to promote statistical literacy worldwide. This project aims to increase the widespread statistical literacy of citizens. It has previously focussed on young people in educational settings which has involved activities such as a statistical literacy competition; however, ISLP has shifted more recently to also consider statistical literacy in adults.

The summary also mentions Statlit.org a website for statistical literacy-related articles, books, and activities. The website is led by Milo Schield, who is a former President of the National Numeracy Network and Director of the W.M. Keck Statistical Literacy Project, based at Augsburg College in the USA. The first goal of both the website and the project is to present statistical literacy as an interdisciplinary activity, overlapping with quantitative reasoning, quantitative literacy, numeracy, and statistical reasoning. The second goal is to present statistical literacy as the study of statisticians in everyday arguments.

Further action identified through the current review includes Eurostat, the statistical office of the European Union (EU). Eurostat's efforts to increase understanding of statistics include training programmes and their guide to European statistics "Statistics Explained" which aims to present statistical topics in a way that is easily understood.

In a less conventional approach, the Italian National Institute of Statistics (Istat) developed a puzzle magazine named "Statistica enigmistica" which uses crosswords and similar puzzles to introduce statistical concepts to non-specialists (Da Valle & Osti, 2016). There seems to be no available evidence on the efficacy of this approach or indication that it was widely distributed.

ProCivicStat is a strategic partnership across six universities and five countries as part of the Erasmus+ Strategic Partnerships. The goal of this project is to promote the inclusion of civil statistics within statistical education. Outcomes of this product include new methods for statistics instruction to help young people understand evidence and statistics about social phenomena. Other outcomes include conceptualising the knowledge requirements for engaging with civic statistics and this research was discussed in the "Definitions" section of this report (ProCivicStat Partners, 2018).

5.3 Summary

Overall, this section of the review identified multiple initiatives, of varied scale, in the statistical literacy space. Some of the initiatives had **little information available** online about their work in statistical literacy, whether they are currently active and their efficacy. From the information available, there appears to be a lot of **variety in these initiatives regarding location, target audience and overarching aims**. Many of the programmes may no longer be active and it may be beneficial to gain further information about these in particular. For example, insight into why these programmes ended, and aspects of them that were successful or unsuccessful could increase the efficacy of future action. Furthermore, **no work was detected during the review process that aimed to connect these different initiatives**. Therefore, this may be a fruitful future endeavour to share learnings across groups.

6 Conclusion

This report includes the findings from a literature review, commissioned by the Office for Statistics Regulation (OSR), to establish the current landscape of statistical literacy research.

The first section of the report covered existing definitions of statistical literacy. There appears to be no consensus on how exactly statistical literacy should be defined but there are some commonly used components including foundational abilities, knowledge of statistical concepts, and critical thinking. An alternative approach for future attempts at defining statistical literacy may be to consider first the context where the definition will be applied and then specify the components required for statistical literacy in that context.

The second section covered evidence on how researchers have attempted to measure components related to statistical literacy. Great variability was observed amongst the general public, in the skills that are linked to statistical literacy. Studies often found that skill level was influenced substantially by demographic factors such as age, gender, and education. No evidence was found that aimed to capture the multidimensional nature of statistical literacy.

The third section outlined evidence on the topic of statistical communication, including research on how to communicate effectively to non-specialist audiences to maximise understanding and engagement. A wide range of recommendations were identified on how to best communicate statistics to non-specialist audiences. The evidence base should be considered when deciding whether to apply these recommendations to statistical communication. The examples of relevant OSR correspondence were largely aligned with the findings of the review although relevant correspondence could not be found for all topics.

Lastly, the review explored actions taken in the UK and internationally to improve statistical literacy and statistical communication. Multiple initiatives, of varied scale, were identified in the statistical literacy space, although some of the initiatives had little information about them available online. It may be beneficial to gain further information about these initiatives to benefit future work. No work was detected during the review process that aimed to connect these different initiatives. Therefore, this may be a fruitful future endeavour to share learnings across groups.

This research was commissioned to shape OSR’s future work on statistical literacy. The findings from the review provide an evidence base that will support OSR in developing its public position on statistical literacy and guide future regulatory work.

Annex A

Approach

The aim of this review was to gain insight into the current landscape of statistical literacy research. At the inception of the review, key terms were used to complete searches in Google and Google Scholar to identify how readily available evidence was in this area. It was established at an early stage that the terms associated with statistical literacy were not sufficiently narrow to filter records to a manageable number to review within the time frame of the review. The search terms were also not leading to a substantial number of highly relevant records.

Therefore, the decision was made to take a more pragmatic approach to the review. This entailed identifying highly relevant records from the initial searches and suggestions by members of the project group at OSR. Once a set of initial relevant articles were collected, their reference lists and papers that cited these articles were then screened in a snowballing approach.

Furthermore, a Twitter scraping tool, developed internally by OSR, was also applied using search terms including: “improving statistical literacy”, “communicating statistics”, “effective communication of statistics” and “general public statistics understanding”. This approach was used to pick up additional recent research in these areas and capture evidence that may not refer to “statistical literacy” specifically. Any relevant evidence detected using this approach had already been picked up in the initial search engine approach.

To gain an additional understanding of the OSR’s previous stance in the areas of statistical literacy that arose in this review, the OSR website was also used to identify any relevant correspondence including key terms. Searches were conducted on published correspondence which are made available on the OSR website.

The initial findings of the review were presented in an OSR-wide session, also attended by Johnny Runge (ESCoE). Attendees were asked for feedback on any potential gaps in the evidence base or misinterpretations of the data. This report also underwent a round of peer-review by researchers identified as key members of the Statistical Literacy community (Johnny Runge and David Spiegelhalter).

Overall, this review and the approach used were intended to provide a broad overview of the literature around statistical literacy rather than to provide a systematic or exhaustive summary of all relevant literature. The results of the review will therefore be limited by the search terms applied and the records identified in the initial stages of the review.

References

Akl, E. A., Oxman, A. D., Herrin, J., Vist, G. E., Terrenato, I., Sperati, F., Costiniuk, C., Blank, D., & Schünemann, H. (2011). Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database of Systematic Reviews*, *3*.

BBC Trust. (2016). *BBC Trust Impartiality Review: Making Sense of Statistics*. http://downloads.bbc.co.uk/bbctrust/assets/files/pdf/our_work/stats_impartiality/report.pdf

Ben-Zvi, D., & Garfield, J. B. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Springer.

Bholat, D., Broughton, N., Parker, A., Ter Meer, J., & Walczak, E. (2018). Enhancing central bank communications with behavioural insights. *Bank of England Working Paper*.

Butt, Swannell, & Pathania. (2022). *Public confidence in official statistics—2021*. https://www.natcen.ac.uk/media/2202247/NatCen_Public-Confidence-in-official-statistics_2021.pdf

Chick, H., & Pierce, R. (2013). The Statistical Literacy Needed to Interpret School Assessment Data. *Mathematics Teacher Education and Development*, *15*(2), n2.

Da Valle, S., & Osti, S. (2016). *"STATISTICA ENIGMISTICA": AN ISTAT PUZZLE MAGAZINE TO INTRODUCE NON-SPECIALISTS TO STATISTICS THROUGH A LUDIC APPROACH*. https://iase-web.org/documents/papers/rt2016/DaValle.pdf?1482484529

Department for Business Innovation & Skills. (2012). *The 2011 Skills for Life Survey: A Survey of Literacy, Numeracy and ICT Levels in England*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/36000/12-p168-2011-skills-for-life-survey.pdf

Department for Business, Innovation & Skills. (2013). *The International Survey of Adult Skills 2012: Adult literacy, numeracy and problem solving skills in England*. https://www.gov.uk/government/publications/international-survey-of-adult-skills-2012

Dhami, M. K. (2018). Towards an evidence-based approach to communicating uncertainty in intelligence analysis. *Intelligence and National Security*, *33*(2), 257–272.

Duffy. (2016). *IPSOS VIEWS. Perils of Perception.* https://www.ipsos.com/sites/default/files/2016-06/015.1_PerilsOfPerception_March2016.pdf

EFSA. (2019). *Guidance on Communication of Uncertainty in Scientific Assessments*. https://www.efsa.europa.eu/en/efsajournal/pub/5520

Ferligoj, A. (2015). How to improve statistical literacy? *Advances in Methodology and Statistics*, *12*(1), 1–10.

Financial Conduct Authority. (2021). *Financial Lives 2020 survey: The impact of coronavirus*. https://www.fca.org.uk/publication/research/financial-lives-survey-2020.pdf

Freeman, A. L., Kerr, J., Recchia, G., Schneider, C. R., Lawrence, A. C., Finikarides, L., Luoni, G., Dryhurst, S., & Spiegelhalter, D. (2021). Communicating personalized risks from COVID-19: Guidelines from an empirical study. *Royal Society Open Science*, *8*(4), 201721.

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A.-S., Broschinski, S., & Politi, M. C. (2012). Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology*, *31*(3), 286.

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

Gal, I. (2004). Statistical literacy. In *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47–78). Springer.

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

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, *31*(3), 444–457.

Galvão, A., Mitchell, J., & Runge, J. (2019). Communicating Data Uncertainty: Experimental Evidence for UK GDP. *Economic Statistics Centre of Excellence Discussion Paper ESCoE DP-2019-20*.

Greis, M., Ohler, T., Henze, N., & Schmidt, A. (2015). Investigating representation alternatives for communicating uncertainty to non-experts. *IFIP Conference on Human-Computer Interaction*, 256–263.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. In *Science* (Vol. 290, Issue 5500, pp. 2261–2262). American Association for the Advancement of Science.

Home Office. (2013). *Statistical and analytical guidance on crime and policing statistics. For analysts working for Police and Crime Commissioners.* https://uksa.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images-statisticalandanalyticalguidanceoncrimeandpolicingstatistic_tcm97-42969.pdf

Ipsos. (2013). *The Perils of Perception*. https://www.ipsos.com/en/perils

Klein, T., Galdin, A., & Mohamedou, E. (2016). An indicator for statistical literacy based on national newspaper archives. *Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE)*.

Kreuzmair, C., Siegrist, M., & Keller, C. (2016). High numerates count icons and low numerates process large areas in pictographs: Results of an eye-tracking study. *Risk Analysis*, *36*(8), 1599–1614.

Lehohla, P. (2002). *Promoting statistical literacy: A South African perspective*.

Mislavsky, R., & Gaertig, C. (2022). Combining probability forecasts: 60% and 60% is 60%, but likely and likely is very likely. *Management Science*, *68*(1), 541–563.

National Numeracy. (2019). *Numerate nation? What the UK thinks about numbers*. https://www.kcl.ac.uk/policy-institute/assets/national-numeracy-day-2019.pdf

OECD. (2012). *Programme for International Student Assessment (PISA) Results from PISA 2012*. https://www.oecd.org/unitedkingdom/PISA-2012-results-UK.pdf

Open Data Institute. (2022). *Data literacy and the UK government*. https://theodi.org/article/data-literacy-and-the-uk-government-report/

Oxman, A. D., Glenton, C., Flottorp, S., Lewin, S., Rosenbaum, S., & Fretheim, A. (2020). Development of a checklist for people communicating evidence-based information about the effects of healthcare interventions: A mixed methods study. *BMJ Open*, *10*(7), e036348.

PARIS21. (2016). *PARTNERING FOR PROGRESS: BUILDING STATISTICAL SYSTEMS THAT GET US TO 2030*. https://paris21.org/sites/default/files/2017-09/P21Board2016_web_0.pdf

PARIS21. (2021). *2021 Update on the indicator for statistical literacy*. Paris 21. https://paris21.org/statistical-literacy-2021

Peters, E., Hart, P. S., & Fraenkel, L. (2011). Informing patients: The influence of numeracy, framing, and format of side effect information on risk perceptions. *Medical Decision Making*, *31*(3), 432–436.

ProCivicStat Partners. (2018). *Engaging Civic Statistics: A Call for Action and Recommendations*. http://IASEweb.org/ISLP/PCS

Runge. (2021). *Communicating Data Uncertainty on GDP and Unemployment: Interviews with the UK Public*. https://escoe-website.s3.amazonaws.com/wp-content/uploads/2021/06/23120020/ESCoE-DP-2021-07.pdf

Runge, J., & Hudson, N. (2020). Public Understanding of Economics and Economic Statistics. *EScoE Occasional Paper*, *3*.

Runge, J., & Killick, A. (2021). *What can Economists Learn from Public Perspectives on the Economy and Economic Statistics? (ESCoE OP-04)*. https://www.escoe.ac.uk/publications/public-perspectives-on-the-economy-and-economic-statistics/

Rutter, H., Parker, S., Stahl-Timmins, W., Noakes, C., Smyth, A., Macbeth, R., Fitzgerald, S., & Freeman, A. L. (2021). Visualising SARS-CoV-2 transmission routes and mitigations. *Bmj*, *375*.

Sabbag, A., Garfield, J., & Zieffler, A. (2018). ASSESSING STATISTICAL LITERACY AND STATISTICAL REASONING: THE REALI INSTRUMENT. *STATISTICS EDUCATION RESEARCH JOURNAL*, *17*(2), 141–160. https://doi.org/10.52041/serj.v17i2.163

Schield, M. (1999). Statistical literacy: Thinking critically about statistics. *Of Significance*, *1*(1), 15–20.

Schield, M. (2006a). *Statistical Literacy Survey Analysis: Reading Graphs and Tables of Rates and Percentages*.

Schield, M. (2006b). Statistical literacy survey results: Reading graphs and tables of rates and percentages. *Conference of the International Association for Social Science Information Service and Technology (IASSIST)*.

Schneider, C. R., Freeman, A. L., Spiegelhalter, D., & van der Linden, S. (2021). The effects of quality of evidence communication on perception of public health information about COVID-19: Two randomised controlled trials. *PloS One*, *16*(11), e0259048.

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, *4*(1), 118–133.

Spiegelhalter, D. (2017). Risk and uncertainty communication. *Annual Review of Statistics and Its Application*, *4*(1), 31–60.

The Plain Numbers Project. (2021). *Plain Numbers Project: Initial trials report*. https://plainnumbers.org.uk/phase-1-project-report

UNECE. (2009). *Making Data Meaningful. Part 2: A guide to presenting statistics*. https://unece.org/DAM/stats/documents/writing/MDM_Part2_English.pdf

UNECE. (2014). *Making Data Meaningful. Part 4: How to improve statistical literacy: A guide for statistical organizations*. https://unece.org/DAM/stats/publications/2013/Making_Data_Meaningful_4.pdf

Van Der Bles, A. M., van der Linden, S., Freeman, A. L., & Spiegelhalter, D. J. (2020). The effects of communicating uncertainty on public trust in facts and numbers. *Proceedings of the National Academy of Sciences*, *117*(14), 7672–7683.

Vicente, M. R., & López, A. J. (2017). Figuring Figures: Exploring Europeans' Knowledge of Official Economic Statistics. *Journal of Official Statistics (JOS)*, *33*(4).

Vicol. (2020). *How to communicate uncertainty*. https://fullfact.org/media/uploads/en-communicating-uncertainty.pdf

Visschers, V. H., Meertens, R. M., Passchier, W. W., & De Vries, N. N. (2009). Probability information in risk communication: A review of the research literature. *Risk Analysis: An International Journal*, *29*(2), 267–287.

Wallman, K. K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, *88*(421), 1–8.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, *2*(2), 3–46.

Ziegler, L. A. (2014). *Reconceptualizing statistical literacy: Developing an assessment for the modern introductory statistics course* [PhD Thesis]. University of Minnesota.

Ziegler, L., & Garfield, J. (2018). DEVELOPING A STATISTICAL LITERACY ASSESSMENT FOR THE MODERN INTRODUCTORY STATISTICS COURSE. *STATISTICS EDUCATION RESEARCH JOURNAL*, *17*(2), 161–178. https://doi.org/10.52041/serj.v17i2.164