International Journal of Communication 16(2022), 2781–2800 1932–8036/20220005

Rethinking the Expertise of Data Journalists: A Case Study

JINGRONG TONG[1]
The University of Sheffield, UK

This article examines the expertise of eight data journalists and its development and application in three British national newspapers and their Sunday counterparts. Apart from technical skills, such as coding techniques, their expertise has a social dimension: the ability to make knowledge/experience-based judgments, solve problems, communicate with audiences, and collaborate with non-data journalists and experts. Their learning of technical skills was initially driven by their personal imperative to stand out in the job market. The social dimension of their expertise builds on and interacts with their gaining and proficient use of technical skills, underlying the combined use of technical and journalistic skills in practice. It helps to achieve collaboration and consolidate cultural authority in telling reliable data stories. The applicability of their expertise transcends all topic areas and suits interdisciplinary, high-tech reporting tasks requiring multifaceted knowledge that a single journalist may not possess. Its application reflects an organizational strategy to modify the division of labor in well-resourced national newspaper newsrooms in the UK in response to the opportunities brought about by the datafication of our society.

Keywords: data journalism, British national newspapers, expertise, cultural authority, the division of labor

Expertise is not merely about skills and knowledge. It is also about performance and interactions with others, in particular, laypeople, which is expertise's social dimension (Goldman, 2018). Traditionally, (beat) journalists have two but partly interacted types of expertise: expertise in subject domains, such as environmental or political issues, and expertise in journalistic practice, such as reporting skills and knowledge to deal with different reporting scenarios and to communicate with audiences (Reich & Godler, 2017). Dealing with difficult reporting situations, communicating with the audience, and finding solutions to problems in the news reporting process are the key journalistic, social skills of "master" journalistic practitioners, particularly those of investigative journalists. These skills are developed and expressed in practice rather than being taught through textbooks (Ettema & Glasser, 1998). The establishment of journalistic expertise underlies the legitimation of journalists' claims to knowledge and truth constructed in news stories (Anderson, 2008).

Jingrong Tong: j.tong@sheffield.ac.uk
Date submitted: 2021-05-24

[1] Thanks go to the journalists for generously sharing their views and the reviewers for their valuable comments.

Data journalism, a form of journalism that involves telling stories in and from data, has become integrated into newsrooms. Examining the expertise of data journalists can offer a useful lens through which to understand data journalism (Lewis & Westlund, 2015). Unlike traditional beat journalists, data journalists usually do not have expertise in specific topics or geographical areas. The current literature sees data journalists' expertise as being skill based, consisting of technical skills such as programming, statistical analysis, data management, visualization, and the ability to use technical tools on top of journalistic methods such as writing (such as Appelgren & Lindén, 2020; Appelgren & Nygren, 2014; Cheruiyot, Baack, & Ferrer-Conill, 2019; Lawson, 2021). Meanwhile, other studies (such as Borges-Rey, 2020; Porlezza & Splendore, 2019) emphasize the importance of collaborations and human interactions with other journalists, particularly beat journalists, and external social actors in data journalists' work. Such collaborations and interactions should involve extensive use of journalistic and social skills. These discussions prompt one to think about whether and to what extent data journalists' expertise has a social dimension along with technical skills, how their expertise is developed and applied in practice, and what implications this has for understanding the role of data journalists' expertise in the newsroom.

To address these questions, this article discusses the expertise of data journalists and its development and application through a case study of eight data journalists in three British national newspapers and their Sunday counterparts. This article begins by discussing the concept of "expertise" and previous expertise studies in multiple disciplines and journalism studies. It then introduces the present study and the methods and data used, followed by a discussion of the nature and composition of data journalists' expertise and its development and application as found in this research. The last section of this article reflects on the implications for our understanding of data journalists' expertise and their role in today's newsrooms.

Conceptualizing "Expertise"

The concept of expertise is well established in multiple disciplines such as sociology, education, psychology, and philosophy. A huge body of studies have examined expertise (such as Abbott, 1988; Chi, Glaser, & Farr, 1988; Collins & Evans, 2007; Farrington-Darby & Wilson, 2006; Glaser, 1985a, 1985b; Johnson, 2013; Nichols, 2017; Norman et al., 2018; Rhodes, Lancaster, & Rosengarten, 2020; Schmidt & Boshuizen, 1993). Drawing from these studies, the article summarizes the following three interrelated aspects of this concept:

1. Expertise refers to skills and knowledge not possessed by others, representing a grasp of a particular subject learned and developed from experience or education (Ericsson, 2014). This is the knowledge dimension of expertise.
2. Apart from the knowledge dimension, expertise also has a social dimension. This dimension refers to experts' performance skills and ability to show others how to solve problems or execute tasks, and their skills and ability should be recognized by others (Goldman, 2018). In other words, an expert needs to use his or her expertise to help nonexperts solve problems. Communication is thus thought to be important to expertise because an expert's expertise must be known and recognized to be meaningful (Treem & Leonardi, 2016).
3. Expertise involves processes such as decision-making and problem-solving processes (Farrington-Darby & Wilson, 2006).

These three aspects point to the importance of three elements to expertise: (1) knowledge and skills, (2) performance and interactions with others, and (3) decision making and problem solving. For professionalism study scholars, expertise helps a profession exert control over its boundaries and establish and maintain autonomy and authority (Abbott, 1988; Larson, 1977).

Expertise has also been examined in journalism studies, although journalism is not a classic profession like medicine or law. Journalism is an occupation of gathering, processing, and producing information and knowledge. Journalists' expertise refers to the "special knowledge and skills" that journalists possess and use to gather and construct news for the public (Anderson, 2008). Traditionally, journalistic expertise is the ability to report reality, comprising knowledge of specific topic areas, the ability to interact with news sources and the audience, and reporting skills (Reich & Godler, 2017). It is also the ability to get a task done faster and better because of the experience (Eyal, 2013).

Expertise in the subject matters of reporting is closely related to news sources and the news beat system. The news beat system, which emerged in the 19th century, fulfils the wish of large news organizations to report news quickly, efficiently, and accurately and boost the productivity of journalists (Manning, 2001). Beat journalists benefit from their expertise in the topic areas they are allocated to cover and their established contacts with experts who can be used as news sources (Fishman, 1980; Magin & Maurer, 2019; Marchetti, 2005). Journalistic labor is divided into and organized around these topic areas in large news organizations (Logan, 2001).

Based on Collins and Evans's (2007) model, Reich (2012) offers a framework for understanding journalistic expertise. He argues that journalists' interactions with (expert) news sources and new audiences in day-to-day news practices help form their interactional expertise. Such expertise includes domain knowledge through reporting and interacting with experts. It also involves the ability to make intuitive judgments about reliable information, news sources, and newsworthiness to produce knowledge about public life. Meanwhile, journalists also need to "translate" (communicate in an easily understood language) abstract knowledge about a particular subject to lay audience members. Their claims to knowledge and truth constructed in their news stories are legitimated by their possession of expertise.

Anderson (2008) sees journalistic expertise as discursive, with the power to influence journalists' control over the boundaries of their jurisdiction and cultural authority in constructing the news. In other words, claims to expertise help journalists consolidate their cultural authority and maintain boundaries (Carlson, 2015). Likewise, Hermida (2015) argues that the ability to assemble and verify facts is an important element of journalistic expertise that can help them claim authority and defend boundaries. Therefore, journalistic expertise has both skill and social dimensions. It refers to reporting skills and domain knowledge. It is also about the ability to demonstrate such skills and knowledge to audience members. The ability to interact with news sources and the audience and make decisions about the truth and accuracy of information and news worthiness is also at the center of journalistic expertise.

Along with the adoption of digital technology in journalism, new forms of journalism and new roles in the newsroom have emerged. One of the new forms of journalism is data journalism. Data journalism is not an entirely new form of journalism, with its history tracing back to the 18th and 19th

centuries (Rogers, 2013). However, modern data journalism was not widely practiced or extensively embedded in newsrooms until around 2008, when leading international news outlets such as *The Guardian* in the UK and *The New York Times* in the United States published compelling data-driven stories.

Over the past 10 years, data journalism has been gradually integrated into newsrooms around the world (Appelgren, Lindén, & van Dalen, 2019; Appelgren & Nygren, 2014; Beiler, Irmer, & Breda, 2020; Cheruiyot et al., 2019; Fink & Anderson, 2015; Tong, 2020; Wright & Doyle, 2018; Zhang & Feng, 2019). Particularly in well-resourced newsrooms in wealthy countries, data teams are launched, and data journalists are hired.

Practicing data journalism, which represents a quantitative oriented form of journalism (Splendore, 2016), data journalists stand out for their technical skills. In addition to their journalistic techniques, data journalists are often seen as possessing computing and data skills that can be applied to any beat—skills that traditional journalists may not have or be willing to learn (Borges-Rey, 2016; Magin & Maurer, 2019). Therefore, data journalists often work as part of a team collaborating with other beats; this fits the description of the establishment of "team-based" systems, facilitating knowledge exchange within the newsroom (Borges-Rey, 2016, 2020; Hermida & Young, 2019; Lewis & Usher, 2014; Nikunen, 2014; Stalph, 2020). This particular work mode of data journalists indicates that technical skills, journalistic techniques, and social skills to interact with other journalists in the newsroom are central to data journalists' work.

It is thus interesting to examine what types of skills data journalists' expertise comprises, what role social skills play, how data journalists develop and combine these skills in their practice, and how to understand the role of data journalists' expertise in the newsroom. These questions also echo Lewis and Westlund's (2015) point about the importance of examining interactions and technical skills to understand journalistic expertise in the context of big data.

The Study

This research addresses the previously mentioned questions by focusing on the context of the UK. It examines the words of eight data journalists in three UK national newspapers (broadsheets) and their Sunday counterparts. These newspapers are anonymized for full anonymity. They will be referred to as Newspaper A (left-leaning), Newspaper B (right-leaning), and Newspaper C (right-leaning). The UK is a good choice because the British news media have been on the frontline of practicing data journalism. *The Guardian*, for example, pioneered data journalism with its compelling data stories, such as the MPs' expenses investigation in 2009 (Rogers, 2009) and the Afghanistan and Iraq war logs in 2010 (Davies & Leigh, 2010; Davies, Steele, & Leigh, 2010).

This study's empirical evidence is mainly from in-depth, semistructured interviews with eight data journalists, including three data editors, at these British national newspapers in 2018 and 2019 (see Table 1). By the time the interviews were conducted, all of the three national newspapers had established their data team comprising three or four reporters, including editors. The eight journalists worked for and played a key role in the data team of their employing newspaper and had established themselves as

renowned data journalists. They constituted more than 60% of the team members of the three newspapers' data teams at the time they were working there. Participant recruitment stopped when the interviews achieved saturation. At the time of the interviews, two participants had changed jobs and were working for news organizations. Therefore, in the interviews, they also discussed their work with their current employers. When analyzing the interviews with them, care was taken to focus on the discussions about their work at the UK newspapers and their general thoughts about data journalism. All in-depth interviews lasted for around one hour and were recorded after gaining the consent of the participants. Although long, intensive qualitative interviews can offer rich data about interviewees' experiences and thoughts (McCracken, 1988), the study has a limited number of interviews, and its findings, therefore, are not generalizable. They only reflect the interview participants' experiences. The interview questions focused on (1) why and how they got into data journalism, (2) how they practiced it, and (3) how they perceived the concept, state, and status of data journalism.

Interview transcripts were uploaded and analyzed in NVivo qualitatively, using thematic analysis that looked for repeated patterns and categories in the text (Braun & Clarke, 2006). The steps proposed by Braun and Clarke (2006) were followed: familiarizing oneself with data through repeated reading of the transcripts, generating initial codes, searching for themes, and reviewing, defining, and naming the themes. The technique of distancing the researcher from the data was used in the analysis process to examine codes and themes with relatively fresh eyes each time the data was examined. The data was fully explored to allow themes to emerge with a particular focus on the following aspects: how data journalists did their job, such as how they collaborated with other journalists; how they collected, processed, and used data for reporting; what expertise was shown and needed for doing the work; and how they developed their expertise. Only those themes with sufficient quotes across transcripts were kept.

Table 1. Overall Information About the Participants.

Interview participant numbers	Newspapers
Three (Participants 1–3)	Newspaper A (Left-leaning)
Two (Participants 4–5)	Newspaper B (Right-leaning)
Three (Participants 6–8)	Newspaper C (Right-leaning)

Findings: Data Journalists' Expertise

Overall, drawing on the interviews, this research reveals that the expertise of the eight data journalists has four parts: (1) technical skills, (2) data knowledge, (3) performance and interactions with others, and (4) decision making and problem solving (see the summary of their expertise in Figure 1). While technical skills are central to their expertise, the knowledge of data, making judgments and decisions about data, and interacting with other people in and outside the newsroom to produce data stories are also essential in their work. Their expertise thus also includes the performance ability and skills to make judgments and solve problems that emerge when using data for reporting. This social dimension of their expertise makes them journalists rather than coders or statisticians. Their proficient technical skills and related data experience and knowledge gained from practice endow them with a high level of performance ability and problem-solving skills. Enabled by technical capacity, their social dimension of expertise—dealing with difficult reporting situations, problem-solving, and verifying data—is akin to the

journalistic and social skills of general journalists, particularly investigative journalists. Moving from the skill and knowledge dimension to the social dimension is equivalent to moving from computational thinking to journalistic logic, which combines technical, journalistic, and social skills in data reporting practices. The social dimension of their expertise is underpinning the integration of data reporting into the day-to-day 24/7 news cycle in the three newsrooms.

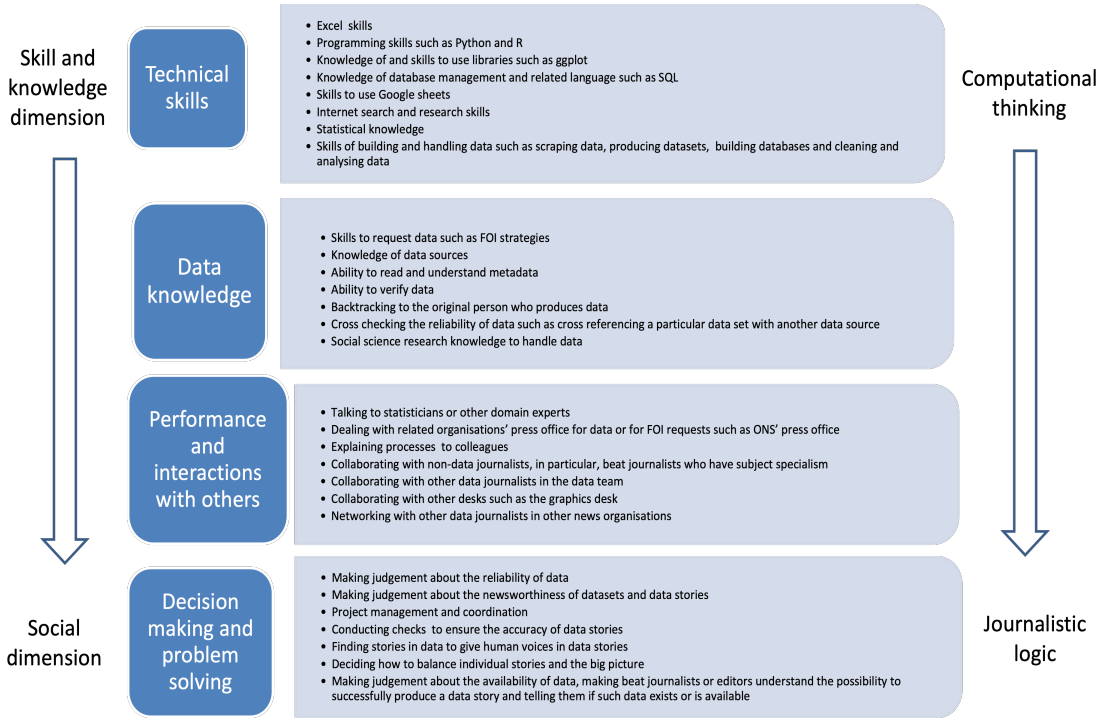

Figure 1. The expertise of the eight data journalists.

Technical Skills

Technical skills refer to the skills of using information technology to practice journalism. Overall, confirming the findings of previous studies (such as Appelgren & Nygren, 2014), all the eight participants possessed technical skills that can be used to collect and handle data and databases, and that non-data journalists may not have. Essential knowledge and skills include programming and statistical analysis skills, knowledge of whether, how, and where they can find the data needed for specific topics, skills in building and handling data, and knowledge about building and managing databases (see more details in Figure 1). Seven (out of the eight) participants had coding skills.

Excel is the essential tool in their toolkit and is usually used to handle small-scale data sets. It is also a tool they used when they were first starting to practice data journalism; Participant 5 said, "When I

started doing it, I would be using an Excel (spreadsheet)."[2] Google Sheets is also seen as a useful tool that data journalists can start with and use to handle smaller data sets. But later, they had to learn a programming language, especially when dealing with large-scale data sets or repetitive work. Learning a programming language and related libraries and packages, such as ggplot and ggplot2, became essential for strengthening their journalism skills.

Python and R were the two dominant programming languages that the participants used in the three data teams at the time of research. Some participants used both, but the rest used one or the other. Apart from Python and R, another programming language, SQL, was also commonly used to handle relational databases. Participant 7 explained the beauty of using programming languages for journalism:

> The great thing about that is to be able to automate something. . . . You do it (write code) once, and then next month, when the data comes out again, you do not have to reinvent the wheel and rewrite the code all over again. You can replicate the code you've already done, so that is one great thing about data journalism and about computing, which is that it enables you to be more efficient.

The choice of programming languages makes a difference in collaboration in the newsroom, especially when it comes to checking each other's code, as shown in Participant 3's remarks: "I have another colleague who writes in Python. We will help each other, and my other colleague has another colleague who practices R so that they can look at each other's code as well." Using the same programming language can also help increase their work efficiency because they can share code or templates, speeding up their data processing.

Most of the time, the participants conducted their own analysis to get the stories, differing from statisticians analyzing data and then supplying findings to journalists. The need to analyze data by themselves extends data journalists' role to include that of data analysts, changing what they should do and their epistemology—how they know what they know. This makes the ability and skills to analyze data part of their expertise.

How Did They Learn and Develop Their Technical Skills?

The participants learned and developed their technical skills mostly by doing it, that is, through experience, self-teaching/learning, and postuniversity learning. All of the eight participants have a degree in journalism or social science and humanities, such as English. Some have a master's degree or even higher in related subjects. Half (of the participants) were introduced to data journalism at university. Getting to know data journalism at university nurtured their interest in it, paving the way for them to secure a job in data journalism and providing a foundation for learning and developing technical skills afterwards. Participant 5 said,

[2] Parentheses in the quotes were all added by the author.

> And I enjoyed that (the data journalism module on the university course). I wanted to do more, and I thought I should try to get into it as a job because it was something where I could sort of mark myself out.

Yet, such skills were mainly learned and picked up from actual practices and postuniversity continuing learning. Participant 4 noted, "So I think you do it just by experience, despite not being a specialist in a particular type of data, you do develop lots of different specialisms just through doing it every year."

Participant 3 also explained how their first data journalism project helped them develop technical skills, understand what data journalism is, and get into data journalism:

> I built a calculator in Excel for all of the different payments, . . . at that point, I did not really identify what I did was data journalism, and it was not a big tradition of it in art. But after that, I started teaching myself the skills properly. And over the course of the next, I do not know how many years, I built up some data journalism skills on them and eventually started. I got my title and then became a data journalist.

Two important points—self-teaching/learning and the help of personal contacts such as peers, friends, and even relatives—emerged from the interviews with all the eight participants. Participant 8 taught themselves programming languages, "from beginners guides to programming," adding, "going to conferences is now quite important for me to my skills because you can learn about different approaches."

While self-teaching, the help from peers, friends, and relatives appears to be important in improving technical skills and speeding up problem-solving. For example, when doing their first data journalism project, Participant 3 gained help from their sibling, who specialized in statistics. While Participant 3 was learning programming languages and ran into difficulty—when "I really cannot understand why something is not working," in the participant's words—they turned to their programmer partner for help. Participant 3 also believed that other data team members who might have more knowledge of statistics were also helpful when they needed such support.

According to Participant 8, peers' support matters because they can pass on experience in problem-solving:

> It is very helpful to be in a team of journalists because, again, your colleagues might have come across similar problems. They can at least point you in the right direction when you need to know how to do something new.

Beyond Technical Skills

Beyond technical skills, other elements of their expertise include performance ability, the ability to interact effectively with others, and the ability to make judgments and decisions about reporting situations to solve problems. A point that arose from all of the eight interviews was that gaining a proper understanding of data and verifying them as much as possible underpin a good data story. That data can

be messy and full of errors makes the ability to make judgments and decisions about the reliability and usefulness of data sets extremely important. Participant 2 explained, "It is really a case of figuring out how useful data sets are, even data sets that have a government agency stamp on them."

To verify data, all the participants needed to have the (journalistic) skills and knowledge to cross-check the data and talk to those close to the data, such as statisticians, and backtrack the data they want to use for reporting. The participants also needed to know how to read metadata—the information about the data—and try to use them to understand the data. Nevertheless, metadata does not always accompany data sets, in which case, the participants had to talk to the people who produced the data and check with them about the data's reliability. For example, Participant 4 explained, "If there is any ambiguity or we are not sure about something (in the data), I just call them up and talk to them, you know, whoever produced data by trying to backtrack through back to the original person."

On the occasions when data did not exist or were not available to them, the participants had to build data sets from scratch through research, interviews, or filing Freedom of Information Act (FOIA) requests. Data verification thus also involved verifying the data sets they built, which required extensive interactions with data holders or the people who were close to the data. Participant 2 gave an example of how to verify the data sets about crime data they created:

> The only way I find to verify so far is to take all of the numbers, to take the details and send them back to the police forces and say: These are all the details that we have so far this year. Can you verify these? Are all of these deaths did take place? If anything you think is missing, please let me know.

Talking to an expert is also crucial for aggregating data without inflating them and for verifying data journalists' data analysis. For example, although they have statistical skills, data journalists are not statisticians. When they use statistical tests or modeling, they must make sure their tests and analysis are correct. Participant 3 commented,

> (For an article that involves the use of statistical tests,) I consulted a friend who is a statistician just to make sure that the tests I was using (were used properly). And then we spoke to some statisticians and one of the other data team about that particular project, . . . just to make sure that the methodology we were using was the right one.

Externally, the participants needed to work closely with data holders and publishers such as the Office for National Statistics and collaborate with experts such as academics for three reasons. First, the data they needed for reporting may not have been open data. Even if they were open data, they may only have been summary data—especially in the UK context, in which personal data protection is prioritized. Second, because data—even government data—may have errors, data journalists need to (or try their best to) verify them before using them for reporting. Third, to appropriately interpret data, data journalists need to understand the context of data collection and production. Detailed (and complete) metadata containing information about data can help them understand the context and interpret the data, yet metadata may not exist. In some cases, interpreting data requires domain knowledge that data

journalists may not have. Therefore, they may need the input from domain experts or data holders and publishers to do the data justice. Given these three reasons, prompt responses from data holders and publishers and experts would be helpful for problem-solving, increasing the importance of data journalists' skills of communicating, networking, and collaborating with them.

Judgments must also be made about the newsworthiness, existence, and accessibility of data sets. In their collaboration with non-data journalists, the participants all experienced a situation in which they had to explain to and convince colleagues why some non-data colleagues' story ideas would not work. The data they requested might not exist, or they might not be able to get them. Participant 4 said, "Sometimes people can come to you with ideas that are not going to work, or the data just does not exist. No matter what we do, we cannot get this data." Likewise, Participant 1 explained that they had to say no to collaboration because they thought "it (was) not doable."

Although having technical skills was key to their work, the participants regarded the ability to tell stories from data as a basic part of being a journalist, which confirms the findings of other studies (see, for example, Lewis & Usher, 2014). The journalistic ability to find human stories that can reflect the bigger picture shown in data—in other words, putting a human face on data—is central to their expertise. The participants primarily relied on their journalistic skills and instincts to identify such human stories and decide which to use in news articles to "translate" data to audiences and communicate with them effectively. Participant 4 gave an example to illustrate how they put a human face on data. In this example, through a community group, Participant 4 found an individual who suffered from the problems the data suggested and was willing to talk about it:

> It (the person's story) was a horrible story, but you know, it is also a perfect representation of putting a human face on that data and being able to get readers to empathize with that story. And if we are just presenting the numbers in themselves, they do not really do the whole job in that particular story.

A last but not least aspect of their expertise is that with more experience gained from intense practices, the participants can ensure the accuracy of their data stories and do the job—revealing truth—faster and better. Their use of data for reporting is continuous, effective, and extensive, differentiating them from non-data journalists who occasionally use data. Participant 3 said,

> I could take the same data set in a much shorter period of time to do all the calculations that I need. With some double checks and verification, I can be sure that this data is solid. . . . I am less likely than my (non-data) colleague to have clarification the following day. The other thing is . . . a question of scale and time. A (non-data) colleague of mine might want to build a data set. It might take them weeks to do it by hand, whereas it might take me days to do programming. So I think it is just the skills. It is the skill set that really makes the difference between (non-data) journalists and data journalists.

The Application of Expertise in Collaboration

The participants' expertise is applied and developed in their collaboration with non-data journalists. In collaboration, they were increasingly "gatekeeping" what to collaborate on, along with their perceived ascending status within the newsroom. Their collaboration with non-data journalists demonstrates their expertise in action, particularly its social dimension enabled by their technical capacity.

Overall, the way that the participants worked was similar across all three national newspapers. All the eight participants were expected to easily switch between beats and collaborate with beat reporters where needed, without being limited to a particular beat, and to establish a good work relationship with them. They often worked in temporarily formed reporting teams to accomplish specific reporting tasks. All the participants embraced collaboration and regarded it as helpful for joining journalists' expertise to strengthen their journalism. For example, Participant 4 commented, "We have subject experts across the newsroom to collaborate with as well to get that knowledge." This quote also illustrates how the participants differentiated themselves from beat reporters: They regarded themselves as journalists with a specialty in data (journalism), whereas beat journalists were the "experts" in subject domains.

Slight organizational variations existed regarding collaboration. Newspaper B assigned data journalists to work with journalists from different desks across the newsroom. Therefore, they could know beat journalists well and familiarize themselves with the topics in various reporting areas. "Embedding with news desks" was seen as essential for data journalists to use their expertise in reporting because they could "then go back to them (beat journalists) and say: this is coming up, should we do something together on it, then we get the benefit of having their knowledge" (Participant 5). Embedding with particular beats for several months made the collaboration between data journalists and beat reporters in Newspaper B more organizationally propelled than in the other two newspapers.

Although collaboration was central to their work, the participants could decide which topics to pursue. They reported different levels of perceived flexibility in choosing their own topics, varying from "50 vs. 50" (Participant 6) to "80 vs. 20" (Participant 2) between the topics suggested by non-data journalists and the topics that data journalists decided to pursue.

When collaborating, the participants did not see themselves as merely playing a supporting role by providing numbers for beat reporters. Participant 2 said, "We are not a service desk." Instead, they wanted to get involved in developing data stories from the very beginning and even fuel beat journalists with story ideas. To explain this point, Participant 2 provided "an example where the data journalists will see something that a regular reporter will not spot and potentially structure and bring the story from their side." In this example, the data reporter found data and made the data (which were in pdf and thus unstructured data, difficult to search and analyze) searchable and analyzable. By so doing, they showed the "ability to bring a story to the (beat) reporter, allowing us to do something different." This example suggests a proactive role that data journalists play in their collaboration with beat journalists. It also shows the expertise of data reporters in knowing where to find data, dealing with data in pdf format, and finding stories in data, as well as how they use their expertise in reporting.

As the preceding discussion shows, the participants needed to decide what to collaborate on, given that they could not collaborate on everything. Because of the time and resource limits, they had to choose to focus on areas in which they had expertise or could develop expertise. They also wanted to work on newsworthy topics with the potential to have a great impact and become exclusive stories. The following quote is exemplary: "Trying to figure out where those resources are best placed to get us exclusive data-driven stories" (Participant 2).

Regarding the choice of reporting topics, the participants also considered their readership and their readers' potential preference. For example, for Newspaper A, "a lot of the stuff that we have done . . . has generally been in projects where there is a big public interest in terms of our readership" (Participant 2). For Newspaper B, our "readers love Brexit stuff, so we do quite a lot of politics articles" (Participant 5). For Newspaper C, the readership was considered relatively older by the participants. Therefore, data stories should balance between not intimating an "older audience" and attracting "younger readers" (Participant 6). Considering their readership, the participants cared about the readability of their data stories and felt the need to have "human voices" in data stories to engage audiences.

Organizational norms and values, the participants' personal interests, and how good or bad their collaborations with beat journalists were all impacted their judgments about what to collaborate and report on. As discussed earlier, the participants would consider newsworthiness and the potential to produce exclusives while making such selections. Another influencing factor identified in the interviews was the level of beat reporters' commitments to such collaborations. Data journalists may collaborate more with those willing to collaborate, which may shape their topic focuses. Besides, data in specific domains may require domain-related knowledge to interpret. The more data journalists worked on data in a domain, the more familiar they became with the data in that particular domain, and the more reports they might produce in that domain.

The participants noted that data journalists were gaining internal recognition and enjoying an ascending status in their news organizations under the social impact created by influential, investigative data stories. Data journalism is considered able to produce exclusive reports that help engage readers and broaden the market shares of news organizations. Data journalists can do what non-data journalists cannot do, and they can be helpful to journalists across the newsroom. Participant 2 said,

> Over time with the project team, we have gained ground because we do have support from other editors within the newsroom to see what we do as important to building stronger stories that allow you to look at a system as a whole rather than individual anecdote, and that is really helpful. . . . we had a couple of stories over time that really kind of helped us to show what we can do, and as reporters work with us . . . (they) get front-page stories, then more people want to work with us.

Along with increasing internal recognition in the newsroom, the participants were given much more freedom over, in Participant 5's words, "the things that we want to pursue rather than being (told to do), I think. It used to be that we would be told what to do."

The preceding discussion suggests that the eight data journalist participants had developed their expertise from years of experience and continue to develop it in their practice of data journalism and collaborations with beat journalists and experts. In addition to technical skills, such as coding techniques and data analysis skills, their expertise also includes knowledge of data and the ability to find human stories in and from data to engage the audience. On top of these is the ability to make judgments and decisions about data reporting, along with the skills of problem-solving and networking with people in and beyond newsrooms. This social dimension of their expertise, which is enabled by their technical capacity, is where their computational thinking meets journalistic logic and where they apply technical skills in journalistic practices. Their accumulated expertise has allowed them to achieve a perceived higher status and recognition in the newsroom.

Discussion and Conclusion: Four Implications

The preceding discussion portrays the four-part typology of the participants' expertise and how it is developed and applied in the newsroom, with four implications for our understanding of data journalists' expertise and its role in the newsroom: (1) The development of data journalists' expertise starts from individual data journalists' initiative, which newsrooms have embraced; (2) data journalists develop their ability and social skills to handle reporting situations in using technical skills to fulfill tasks and in collaboration with others; (3) data journalists' expertise helps legitimate their claims to knowledge and the truth constructed in data stories; and (4) data journalists' expertise is used to mitigate the traditional news beat system's rigidity, helping to get the full benefits of institutional resources.

First is the importance of data journalists' initiative in gaining expertise. The participants' journey in data journalism began with their initiative to adapt to the job market's needs. Against the backdrop of job cuts and newsroom closure, landing a decent job in a prestigious newsroom is not common among young journalists. Their ability to use technical skills to handle data for journalism enabled them to stand out and succeed in the competitive job market. Most of the participants earned their place and rose through ranks quickly in these esteemed national newspapers as young reporters, some even in their 20s. Their success demonstrates the crucial role of their initiative, determination, and self-teaching and learning in developing their expertise and initiating the rise of data journalism.

These three national newspapers are among the first group of news media in the world that have opted to practice data journalism systematically. They identified the opportunity offered by data to renovate their journalism and thus embraced individual data journalists' endeavors. Datafication is an omnipresent process of transforming our social lives and actions into quantified data (Mayer-Schönberger & Cukier, 2013). The global open data movement was accelerated by governments' open data initiatives, starting with Barack Obama signing his Memorandum on Transparency and Open Government to boost the government's transparency level in 2009. The datafication of our societies has generated opportunities for journalism to be renovated, hold power accountable, and revive journalistic legitimacy (Hermida & Young, 2019; Tong, 2018).

Data journalism is thought to be able to produce influential exclusives and boost subscriptions. The proliferation of data offers huge resources that journalists can deploy. Data stories can be told in novel, interactive ways to engage the audience. Data journalism has the potential to produce compelling

front-page investigative stories that can boost news media subscriptions. All of these possibilities are attractive to the three newspapers, which have been suffering revenue losses since the turn of the 21st century and are try hard to engage audiences and boost their income.

However, seizing such opportunities requires journalists to master related skills and knowledge. Given beat journalists' expertise in their reporting areas, their practicing data journalism would be ideal. Yet, newsrooms would need to invest hugely to upskill the entire workforce to achieve this, which is financially unrealistic (Tong, 2020). Thus, launching a data team comprising data journalists who have developed expertise that can be used for all beats would be a perfect solution for embracing the opportunities brought about in our datafied societies without enormous financial costs.

The second is that the ability and social skills to handle data-related reporting situations are at the heart of the participants' expertise and developed in actual performance and interactions with people. Their expertise was established and developed as they used technical skills to fulfill tasks and collaborate with others in the newsroom and beyond. This study confirms the findings of previous studies (such as Borges-Rey, 2020; Hermida & Young, 2017) that collaborations are at the center of data journalists' work. Their collaborations with external experts are similar to those between data journalists and social actors discussed in the study of Italian data journalism (Porlezza & Splendore, 2019). However, unlike the Italian study, the present study does not present strong evidence of the participants collaborating with data journalists from other newsrooms to conduct data investigations and produce data stories. Apart from some unusual occasions, they mostly collaborated with beat journalists in the newsroom. Their networking with other data journalists from different news organizations was mainly about learning new technical skills and problem-solving knowledge. In extensive collaborations with other journalists in the newsroom and external experts, the participants gained experience and expertise to make judgments about data-related reporting situations and solve related problems. With sufficient experience, they learned to combine technical, journalistic, and social skills in data reporting to the point where their decision making became journalistically intuitive. This social dimension, enabled by their technical capacity and social skills, facilitated their increasing status within the newsroom and sped up the integration of data reporting into daily reporting.

The organizational need to use resources effectively and make the best of journalists' expertise is driving the collaboration within the newsroom. The idea is to put together and mix all kinds of expertise in a big knowledge pool and flexibly take out and combine specific expertise when needed. This practice confirms other scholars' findings about the emergence of knowledge exchange within news organizations (see, for example, Nikunen, 2014; Reich & Godler, 2017). However, differing from their arguments, this study does not see the rise of data reporting as indicating the entire end of "compartmentalization" in the three UK newspaper newsrooms. The end of compartmentalization means that the division of labor in the newsroom stops being organized around subject areas and appearing along with the traditional news beat system (Logan, 2001). In these three newsrooms, news beats still hold their ground, and their expertise and values stand. The rise of data journalists' expertise opposes the previously recognized tendency toward shrinking specialization in the digital age (Bagley, 2013). The gaining and development of data journalists' expertise benefit from newsrooms establishing a "team-based system" to draw on multiskilled

knowledge to handle interdisciplinary, high-tech reporting tasks. This system differs from the traditional way of beat reporting in which beat journalists primarily work individually.

The need to collaborate and combine different specialisms in news reporting transforms the socialization of journalists, lessening the individualization of news work and increasing the demand for extensive group work. The collaboration between data journalists and beat journalists requires them to commit to collective reporting tasks and to recognize each side's contribution and specialism. Constructing a strong sense of community and belonging among data and beat journalists is much needed for increasing group cohesion. In addition, the need to collaborate stresses the importance of data journalists being present in the office, sitting down with other journalists, and going through projects. This need may mean the return of physical, social space that has allegedly diminished as journalists have turned to the Internet for news sources (Broersma & Graham, 2012; Dick, 2012). Likewise, data journalists need to construct and maintain good work relationships with external experts. In so doing, data journalists can get to know non-data colleagues and experts, and vice versa. It is also essential for data journalists to consider how to communicate with audiences. Putting human voices in data stories to increase their readability can help audiences accept the new form of data journalism.

Third, the participants' expertise, especially problem-solving skills and the ability to make judgments about data reliability, verify data by using technical skills and data knowledge, as well as collaborate with experts, helps them differentiate themselves from non-data journalists and consolidate cultural authority in producing trustworthy stories from data. The participants' words convey a strong sense of authority seeking and defending. This echoes Borges-Rey's (2016) argument that UK data journalists constructed a performative discourse, projecting legitimate performativity to audiences. Where their knowledge might be insufficient to do this properly, the participants turned to use experts' authority to endorse and strengthen their claims to reliable knowledge and the truth constructed in data stories.

Finally, the use of data journalists' expertise in collaborative reporting has an implication for the division of labor in the newsroom. Their expertise is functionally "flexible" and applicable to any beat, which helps to overcome the rigidity of the traditional news beat system and optimize news organizations' resource investments. At the time of research, data journalism was treated as a specialism rather than a normal practice. It was separated from the traditional news beat system, although non-data journalists may also be able to pick up and learn skills to practice data journalism, and data journalists' specialism may have different focuses—for example, some might have a specialty in medical data, while others specialize in environmental data. This separation reorganizes and optimizes resources and labor in the newsroom so that data journalists' expertise connects with that of beat journalists to fulfill specific complex, interdisciplinary reporting tasks. Even if data reporting did eventually become a normal practice, data journalists would still have the expertise to do the work better and faster than non-data journalists. Data journalists' expertise, together with the increasing internal recognition and their compelling data stories, helps them establish cultural authority. These changes in the division of labor in well-resourced newsrooms may enlarge the gap between well and poorly resourced newsrooms in terms of their ability to cover complicated, interdisciplinary topics. Newsrooms with fewer resources may not have the luxury of hiring a team of data specialists to suit the interdisciplinary reporting need in our datafied society.

The present study can contribute to our understanding of data journalists' expertise. The study, however, is limited by its focus on only these British national newspapers and by a small number of interviews. For future research, it would be helpful to examine whether data journalists in other news outlets, such as regional and broadcast news media in the UK and beyond, also possess the participants' expertise. It would also be worth examining whether and to what extent such expertise resembles that of other skill-based forms of journalism, such as mobile journalism and video journalism. In addition, there has been a dramatic increase in the importance of data journalism during the COVID-19 pandemic. It is crucial to research the impact of the pandemic on the expertise of data journalists and its development and application in the newsroom.

References

Abbott, A. (1988). *The system of the professions. An essay on the division of expert labor*. Chicago, IL: University of Chicago Press.

Anderson, C. (2008). Journalism: Expertise, authority, and power in democratic life. In D. Hesmondhalgh & J. Toynbee (Eds.), *The media and social theory* (pp. 262–278). London, UK: Routledge.

Appelgren, E., & Lindén, C.-G. (2020). Data journalism as a service: Digital native data journalism expertise and product development. *Media and Communication, 8*(2), 62–72. doi:10.17645/mac.v8i2.2757

Appelgren, E., Lindén, C.-G., & van Dalen, A. (2019). Data journalism research: Studying a maturing field across journalistic cultures, media markets and political environments. *Digital Journalism, 7*(9), 1191–1199. doi:10.1080/21670811.2019.1685899

Appelgren, E., & Nygren, G. (2014). Data journalism in Sweden: Introducing new methods and genres of journalism into "old" organizations. *Digital Journalism, 2*(3), 394–405. doi:10.1080/21670811.2014.884344

Bagley, K. (2013, January 17). About a dozen environment reporters left at top 5 U.S. papers. *Inside Climate News*. Retrieved from https://insideclimatenews.org/news/20130114/new-york-times-dismantles-environmental-desk-climate-change-global-warming-journalism-newspapers-hurricane-sandy

Beiler, M., Irmer, F., & Breda, A. (2020). Data journalism at German newspapers and public broadcasters: A quantitative survey of structures, contents and perceptions. *Journalism Studies, 21*(11), 1571–1589. doi:10.1080/1461670X.2020.1772855

Borges-Rey, E. (2016). Unravelling data journalism. *Journalism Practice, 10*(7), 833–843. doi:10.1080/17512786.2016.1159921

Borges-Rey, E. (2020). Towards an epistemology of data journalism in the devolved nations of the United Kingdom: Changes and continuities in materiality, performativity and reflexivity. *Journalism, 21*(7), 915–932. doi:10.1177/1464884917693864

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. doi:10.1191/1478088706qp063oa

Broersma, M., & Graham, T. (2012). Social media as beat: Tweets as a news source during the 2010 British and Dutch elections. *Journalism Practice, 6*(3), 403–419. doi:10.1080/17512786.2012.663626

Carlson, M. (2015). Introduction: The many boundaries of journalism. In M. Carlson (Ed.), *Boundaries of journalism* (pp. 1–18). London, UK: Routledge.

Cheruiyot, D., Baack, S., & Ferrer-Conill, R. (2019). Data journalism beyond legacy media: The case of African and European civic technology organizations. *Digital Journalism, 7*(9), 1215–1229. doi:10.1080/21670811.2019.1591166

Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Collins, H., & Evans, R. (2007). *Rethinking expertise*. Chicago, IL: University of Chicago Press.

Davies, N., & Leigh, D. (2010, July 25). Afghanistan war logs: Massive leak of secret files exposes truth of occupation. *The Guardian*. Retrieved from https://www.theguardian.com/world/2010/jul/25/afghanistan-war-logs-military-leaks

Davies, N., Steele, J., & Leigh, D. (2010, October 22). Iraq war logs: Secret files show how U.S. ignored torture. *The Guardian*. Retrieved from https://www.theguardian.com/world/2010/oct/22/iraq-war-logs-military-leaks

Dick, M. (2012). The re-birth of the "beat": A hyperlocal online newsgathering model. *Journalism Practice, 6*(5–6), 754–765. doi:10.1080/17512786.2012.667279

Ettema, J., & Glasser, T. (1998). *Custodians of conscience: Investigative journalism and public virtue*. New York, NY: Columbia University Press.

Ericsson, K. A. (2014). Expertise. *Current Biology, 24*(11), R508–R510. doi:10.1016/j.cub.2014.04.013

Eyal, G. (2013). For a sociology of expertise: The social origins of the autism epidemic. *American Journal of Sociology, 118*(4), 863–907. doi:10.1086/668448

Farrington-Darby, T., & Wilson., J. R. (2006). The nature of expertise: A review. *Applied Ergonomics, 37*(1), 17–32. doi:10.1016/j.apergo.2005.09.001

Fink, K., & Anderson, C. W. (2015). Data journalism in the United States. *Journalism Studies, 16*(4), 467–481. doi:10.1080/1461670X.2014.939852

Fishman, M. (1980). *Manufacturing the news*. Austin, TX: University of Texas Press.

Glaser, R. (1985a). *The nature of expertise*. Columbus, OH: National Center for Research in Vocational Education.

Glaser, R. (1985b). *Thoughts on expertise*. Pittsburgh, PA: University of Pittsburgh Learning Research and Development Center. Retrieved from https://files.eric.ed.gov/fulltext/ED264301.pdf

Goldman, A. I. (2018). Expertise. *Topoi, 37*(1), 3–10. doi:10.1007/s11245-016-9410-3

Hermida, A. (2015). Nothing but the truth: Redrafting the journalistic boundary of verification. In M. Carlson (Ed.), *Boundaries of journalism* (pp. 37–50). London, UK: Routledge.

Hermida, A., & Young, M. L. (2017). Finding the data unicorn. *Digital Journalism, 5*(2), 159–176. doi:10.1080/21670811.2016.1162663

Hermida, A., & Young, M. L. (2019). *Data journalism and the regeneration of news*. London, UK: Routledge.

Johnson, T. (2013). Expertise and the state. In M. Gane & T. Johnson (Eds.), *Foucault's new domains* (pp. 147–160). London, UK: Routledge.

Larson, M. S. (1977). *The rise of professionalism: A sociological analysis*. Berkeley: University of California Press.

Lawson, B. T. (2021). Hiding behind databases, institutions and actors: How journalists use statistics in reporting humanitarian crises. *Journalism Practice*. Advance online publication. https://doi.org/10.1080/17512786.2021.1930106

Lewis, S. C., & Usher, N. (2014). Code, collaboration, and the future of journalism. *Digital Journalism, 2*(3), 383–393. doi:10.1080/21670811.2014.895504

Lewis, S. C., & Westlund, O. (2015). Big data and journalism. *Digital Journalism, 3*(3), 447–466. doi:10.1080/21670811.2014.976418

Logan, R. A. (2001). News' compartmentalization: Implications for food biotechnology coverage. *AgBioForum, 4*(3&4), 194–198.

Magin, M., & Maurer, P. (2019). Beat journalism and reporting. In *Oxford research encyclopedia of communication.* Oxford, UK: Oxford University Press. doi:10.1093/acrefore/9780190228613.013.905

Manning, P. (2001). *News and news sources: A critical introduction*. London, UK: SAGE Publications.

Marchetti, D. (2005). Subfields of specialised journalism. In R. Benson & E. Neveu (Eds.), *Bourdieu and the journalistic field* (pp. 64–82). Cambridge, UK: Polity.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. London, UK: Murray.

McCracken, G. (1988). *The long interview*. London, UK: SAGE Publications.

Nichols, T. (2017). *The death of expertise: The campaign against established knowledge and why it matters*. Oxford, UK: Oxford University Press.

Nikunen, K. (2014). Losing my profession: Age, experience and expertise in the changing newsrooms. *Journalism, 15*(7), 868–888. doi:10.1177/1464884913508610

Norman, G. R., Grierson, L. E., Sherbino, J., Hamstra, S. J., Schmidt, H. G., & Mamede, S. (2018). Expertise in medicine and surgery. In G. R. Norman, L. E. M. Grierson, J. Sherbino, S. J. Hamstra, H. G. Schmidt, & S. Mamede (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 331–355). Cambridge, UK: Cambridge University Press.

Porlezza, C., & Splendore, S. (2019). From open journalism to closed data: Data journalism in Italy. *Digital Journalism, 7*(9), 1230–1252. doi:10.1080/21670811.2019.1657778

Reich, Z. (2012). Journalism as bipolar interactional expertise. *Communication Theory, 22*(4), 339–358. doi:10.1111/j.1468-2885.2012.01411.x

Reich, Z., & Godler, Y. (2017). The disruption of journalistic expertise. In C. Peters & M. Broersma (Eds.), *Rethinking journalism again: Societal role and public relevance in a digital age* (pp. 76–92). London, UK: Routledge.

Rhodes, T., Lancaster, K., & Rosengarten, M. (2020). A model society: Maths, models and expertise in viral outbreaks. *Critical Public Health, 30*(3), 253–256. doi:10.1080/09581596.2020.1748310

Rogers, S. (2009, June 18). MPs' expenses: All the revelations, as a spreadsheet. *The Guardian*. Retrieved from https://www.theguardian.com/news/datablog/2009/may/13/mps-expenses-houseofcommons

Rogers, S. (2013, March 15). John Snow's data journalism: The cholera map that changed the world. *The Guardian*. Retrieved from https://www.theguardian.com/news/datablog/2013/mar/15/john-snow-cholera-map

Schmidt, H. G., & Boshuizen, H. P. (1993). On acquiring expertise in medicine. *Educational Psychology Review, 5*(3), 205–221. doi:10.1007/BF01323044

Splendore, S. (2016). Quantitatively oriented forms of journalism and their epistemology. *Sociology Compass, 10*(5), 343–352. doi:10.1111/soc4.12366

Stalph, F. (2020). Evolving data teams: Tensions between organizational structure and professional subculture. *Big Data & Society, 7*(1). doi:10.1177/2053951720919964

Tong, J. (2018). Journalistic legitimacy revisited: Collapse or revival in the digital age? *Digital Journalism, 6*(2), 256–273. doi:10.1080/21670811.2017.1360785

Tong, J. (2020). Paradigm reinforcing: The assimilation of data journalism in the U.K. *Journal of Applied Journalism and Media Studies*. Advance online publication. doi:10.1386/ajms_00043_1

Treem, J. W., & Leonardi, P. M. (2016). What is expertise? Who is an expert? Some definitive answers. In J. W. Treem & P. M. Leonardi (Eds.), *Expertise, communication, and organizing* (pp. 1–24). Oxford, UK: Oxford University Press.

Wright, S., & Doyle, K. (2018). The evolution of data journalism: A case study of Australia. *Journalism Studies, 20*(13), 1811–1827. doi:10.1080/1461670X.2018.1539343

Zhang, S., & Feng, J. (2019). A step forward? Exploring the diffusion of data journalism as journalistic innovations in China. *Journalism Studies, 20*(9), 1281–1300. doi:10.1080/1461670X.2018.1513814