methods, data, analyses | Vol. 19(2), 2025, pp. 46-86 DOI: 10.12758/mda.2024.07

Preferences, Participation, and Evaluation of Answering Questions About the Books Participants Have at Home Through Conventional and Image-Based Formats

Patricia A. Iglesias

Research and Expertise Centre for Survey Methodology (RECSM), Universitat Pompeu Fabra

Abstract

The collection of photos through online surveys has emerged as a valuable research tool given the growing use of smartphones, which have facilitated the capture and share of photos. However, gaps persist in understanding respondents' involvement in these tasks when asked to perform them in an online survey. Existing literature lacks insights into participants' preferences, their assessment of questions asking for photos, and how their characteristics might impact their participation in such queries. This paper addresses these gaps, while also comparing how image-based formats compare to conventional ones. Conducted among 1,270 parents living with children in primary school of an opt-in panel in Spain, the mobile online survey implemented in this study revealed a preference for conventional questions, and higher participation in that format than in the image-based one. Respondents able to choose their response format and preferring images presented higher participation rates than those without a choice. While both formats were perceived as equally easy, participants using conventional formats liked the questions better than those answering through photos. Finally, age, being female, having a tertiary education degree, and using the camera at least once a week positively impacted the participation in image-based questions, whereas comfort with new technologies increased the likelihood of liking this format. This study not only fills critical gaps in the literature but also sheds light on the complexities of asking for photos in online surveys.

Keywords: image collection, mobile online surveys, books at home, participation, respondent's evaluation, preferred response format

Although surveys are one of the most used methods to collect data, they suffer from measurement errors (Saris & Gallhofer, 2014). The increasing use of smartphones provides new measurement opportunities that could help reducing such errors (Revilla, 2022). In particular, smartphones have sensors, like the GPS, that allows tracking respondents' location and offer in-the-moment surveys (Ochoa, 2022) or the microphone, that can be used to record voice answers (Höhne & Gavras, 2022). Photos captured with the smartphone camera have also gained attention as a possible new data source, since they are expected to alleviate respondents' burden, enhance data accuracy and quality, and provide insights beyond what conventional response formats can offer (Revilla, 2022).

Research on the feasibility of requesting for photos in online surveys has focused mainly on the respondents' willingness to capture and share photos. Further, studies assessing participation in such questions have asked mostly for general photos (e.g., of the respondents' surroundings, see Bosch et al., 2019), but there is little evidence for more demanding tasks (e.g., submitting multiple photos or capturing items in various locations). Furthermore, scant attention has been devoted to investigating respondents' preferences, or to exploring their evaluation of such response formats. This lack of knowledge does not allow practitioners to make informed decisions regarding the use of visual data within the frame of online surveys: even though images could increase data quality, their relevance diminishes if respondents do not provide such photos.

Additionally, certain participants might be more inclined to participate in image-based response formats than others, potentially introducing biases into who submits photos and who does not. Thus, in this paper I provide new evidence about the respondents' preferences, levels of participation, and evaluation of questions asking to share visual data focusing on a more demanding task: capturing photos of all the books respondents have at home. Since inquiries about the number of books at home have been a recurring feature in numerous surveys within this field, this case study can both enhance our understanding of the efficacy of collecting visual data and facilitate substantive analyses in the realm of social sciences.

Traditionally, the number of books at home has been used to measure cultural capital and/or socioeconomic status (see Heppt et al., 2022; Sieben & Lechner, 2019). Further, analyses on this question show that the number of books impacts dimensions such as parent health literacy (Sanders et al., 2004), socio-emotional skills (Brunello et al., 2012), reading scores (Güre et al., 2023; McNally et al., 2023), and students' academic language comprehension (Heppt et al., 2022).

Direct correspondence to
Patricia A. Iglesias, Research and Expertise Centre for Survey Methodology (RECSM), Department of Political and Social Sciences, Universitat Pompeu Fabra, C/ de Ramon Trias Fargas, 25-27, 08005, Barcelona, Spain
E-mail: patricia.iglesias@upf.edu

However, this question exhibits limitations. First, respondents might answer the survey from a place different than home, preventing them from looking at the books they have, leading to inaccurate answers. Second, even if respondents are at home during the survey, counting each book can be arduous, especially if they have many. Thus, this question would mostly rely on estimates, potentially compromising accuracy. Third, the previous issues seem to have been addressed by presenting response intervals corresponding to the number of books. However, the large intervals affect the granularity. For instance, Gil Flores (2011) used the categories 0-10, 11-25, 26-100, 101-200, and more than 200 books. These categories do not provide detailed information on the exact number of books; having 26 books is very different from having 100 books. Fourth, using intervals could increase social desirability bias (having more books is well-regarded), since respondents might avoid selecting categories that fall in the lower range, which can cause an overestimation of the number of books at home.

One might argue that both the lack of granularity and possible social desirability bias could be mitigated by proposing an open-ended question (wherein respondents should type in the exact number of books). This approach might enhance the level of detail in the information and prevent respondents from inferring what is considered a low or high number of books. However, little research has asked for the number of books in this manner, and comparisons of the quality of both estimates have not been conducted to the best of my knowledge, let alone with recent data.

Moreover, count variables also have limitations. For instance, we can expect rounding errors, estimation errors due to satisficing and/or to low mathematical/spatial abilities, and a tendency to over-report the number of books. These problems could be reduced by measuring the number of books in a different way: through photos of the books sent by respondents. Photos of the books could transcend the wide intervals commonly found in questionnaires by offering a count directly observable in the images. Furthermore, photos can yield novel insights, such as the title of some of the books, or the methods employed for storage. Surveys including the question on the number of books usually do not explore such aspects. Knowledge on this could enrich the understanding of cultural and/or economic capital, since researchers might achieve better characterizations of their subjects. For instance, 40 books of English literature might relate differently to cultural capital than 40 cooking books. The language in which books are written could also indicate that participants are familiar with languages different than their native tongue, possibly also expanding their cultural capital. Further, storing books inside boxes or closets might affect how individuals in the dwelling engage with those books, by making them harder to reach. Thus, besides the number of books at home, I also ask for the languages of the books and their storage methods.

This paper investigates survey participants' preferences for, participation in, and evaluation of answering questions in a mobile online survey about the books they have at their main residence using two response formats: a conventional one (radio buttons and textboxes) and an image-based one (taking and sharing photos of the books). In addition, it examines the impact of respondent characteristics on their participation in and evaluation of each format. Thus, the core aims are two-fold: first, to shed light on the feasibility of using conventional versus image-based formats for collecting information; and second, to identify which individuals participate and positively evaluate each format when a more complex task is involved: providing information about the books they have access to at home. The analysis is conducted using data from the Netquest opt-in online panel in Spain collected in June 2023 among parents that had at least one child in primary school. While the quality of the data provided is a further consideration relevant to the evaluation of the image-based format, it falls beyond the scope of this paper.[1]

Background

The research presented in this section regards visual data produced during the survey, in line with the type of data asked in this paper. While there are also studies about sharing visual data already captured (i.e., produced before the survey), they are not explored in this section.

Respondents' Preferences, Participation, and Evaluation

Previous research has studied the feasibility of collecting visual data through online surveys by investigating two main aspects: the respondents' stated willingness to share visual data, and their actual sharing.

Regarding the willingness of respondents to share photos or videos, different results have been reached depending on the type of visual data file and the topic covered. For instance, 56% of participants in an online survey performed in the Netquest panel in Spain would accept to answer questions by taking and sharing photos of products with their smartphone (Revilla et al., 2019), whereas 65% of participants in wave 9 of the Understanding Society Innovation Panel in the UK would use their smartphone camera to take photos or scan barcodes for a survey (Wenz et al., 2019). Thus, the willingness is moderately high (over 50%) when it comes to photos of objects within the household or receipts. However, willing-

[1] Due to the complexity of evaluating data quality, especially in the case of images (this aspect requires a lot of additional analyses and explanations), it will be discussed in a separate paper.

ness seems to be lower for inquiries that might be considered as more intrusive: 18% of respondents in a survey using the online probability-based LISS panel in the Netherlands would be willing to share a selfie (Struminskaya, Toepoel, et al., 2021), while 14% of participants in a survey from Statistics Netherland stated willingness for the same task (Struminskaya, Lugtig, et al., 2021). Regarding a photo of their house, 38% in the first survey would be willing and only 12% in the second survey. Finally, 24% and 16%, respectively, would send a video of their surroundings.

Iglesias and Revilla (2024) asked for the willingness to capture and share a photo with the smartphone among participants of the Netquest opt-in online panel in Spain. 62% of respondents expressed willingness, while 31% stated it depended on the photo they are asked. This hints that the willingness for capturing and sharing photos is high: if put together, almost 93% of respondents would be willing to send photos. This study isolated the potential effect of skills and availability of producing visual data during the survey by asking respondents to not consider those aspects when answering. Moreover, the authors found that 99% of participants stated knowing how to capture photos with the camera in their mobile device, and that 83% would be able to take a photo of something in their house. By analyzing these three dimensions together, the authors estimated the expected participation, and found that 54% of respondents answering from home would participate in an online survey question asking them to capture and share a photo of something at their dwelling.

Additionally, a second set of studies have asked respondents to answer questions by actually sharing images captured during the survey. Some of these studies asked for screenshots: in particular, Ohme et al. (2021) and Sewall et al. (2022) requested screenshots of the iOS Screen Time function. The participation rate was 12% for the first study, and 78% in the last wave of the second study. They used different samples: Ohme et al. (2021) recruited a sample with an opinion research company in the Netherlands, while Sewall et al. (2022) used participants from the Prolific online panel in the United States with a task-approval of 95% or higher.

A study among Millennials from the Netquest panel in Spain and Mexico asked for a photo of the surroundings. 49% of respondents in Spain and 57% in Mexico sent such photos, whereas 24% and 17% skipped the question, respectively (Bosch et al., 2019). This proves to be higher than the willingness stated by survey respondents in the Netherlands regarding the surroundings, which might be explained by them being asked for a video rather than photos. Certainly, capturing a photo allows more control than a video (a photo can be quickly recaptured, while a video could be more burdensome).

Further, Bosch et al. (2022) used the Respondi panel in Germany to ask for photos captured during the survey (with a smartphone) related to given topics, and their equivalent in conventional format. The authors found that breakoff

was higher for image requests in almost all cases, and that, among those not breaking off, participation was significantly greater in the group answering in conventional ways (over 99%) than in the groups doing so through image-based response formats (49% to 67%). When asked for their evaluation, respondents reported liking it better and finding the conventional response format easier than the image-based one.

Ilic et al. (2022) carried out an experiment with participants of the LISS panel, comparing conventional and image-based response formats. In addition to assigning participants to one of these two formats, a third group was introduced, in which respondents could choose their preferred response format for providing information. The authors asked conventional questions or photos of the respondents' favorite place in their home, an outdoor space of their dwelling (e.g., balcony), and their heating system. 57% of participants who had the option to choose their response format preferred answering through photos. Compliance rates[2] varied between the response formats, from 58% to 99% for conventional response formats and from 27% to 78% for image-based ones. Among participants with a choice, those who opted for image-based responses demonstrated higher compliance rates (from 50% to 78%) than those who were solely instructed to send photos (27% to 39%). Differences between the two groups answering through conventional formats were less pronounced (with a maximum difference of 9 percentage points). Non-complying respondents answering questions about their favorite place and heating system with images were asked for their reasons to not capture and submit photos. Half of participants stated privacy concerns regarding the question for their favorite place, and 10% reported technical problems. As for the heating system, 49% stated it was unreachable and 15% mentioned privacy concerns.

The results from these studies show that both willingness and participation in survey questions asking for photos not only varies among, but also within studies when asked for different types of images. This could be associated with the content of the photo that is being asked (e.g., more or less personal/sensitive, more or less difficult/burdensome to capture it), but also with the exact way in which it was asked, different levels of incentives, differences regarding the type of panel or the target populations (e.g., different countries, age cohorts, etc.), as well as temporal changes. These differences in the findings make it necessary to continue exploring other topics and types of tasks. More research is needed to understand the extent to which previous results can be generalized to different situations.

Further, participation in image-based response formats is only contrasted with their equivalent in conventional response formats in few studies, leading to limited evidence about the performance of both formats in similar settings.

---

[2] Measured as providing an answer for the conventional response format, and submitting a photo for the image-based response format.

Finally, respondents' preferences and evaluations have been even less studied so far: to the best of my knowledge, they have been reported by only one study each.

Impact of the Respondents' Characteristics

Previous research has studied the effects of the respondents' sociodemographic characteristics, experience as panelist, and their use of mobile devices on their willingness to share images in the frame of surveys, as well as their expected and actual participation when proposed image-based response formats. Summary tables of the variables impacting these dimensions are available in the supplementary material 1.

As for willingness, Struminskaya, Toepoel, et al. (2021) found no effects for gender and education, but the frequency of taking photos, trust in the anonymity of the answers, and being older than 65 years old positively impacted the likelihood of being willing to share sensor-collected data (including, among others, capturing a photo of the house, of self, and a video of the surroundings). Conversely, Iglesias and Revilla (2024) found a negative relation between age and willingness, but similarly to the previous study, they found no effects for gender and education. Finally, higher participation in previous surveys positively impacted the willingness to participate in questions asking for photos.

Concerning the expected participation, Iglesias and Revilla (2024) found that it is less likely for older respondents to take and share photos of something inside their dwelling. In contrast, the higher the number of surveys completed in the three months prior to the study, the more respondents are expected to share such photos. Gender and education did not impact the expected participation in this study.

Finally, Struminskaya, Lugtig et al. (2021), focusing on actual participation,[3] found that the frequency of use of the camera did not have an impact, and neither did the level of education or participation in previous surveys. Moreover, age impacted positively sharing a photo of the house or of the respondents, and females were more likely to share photos of receipts.

Overall, there is consensus on the lack of impact of education, while either no or positive effects are found for gender, frequency of taking photos, and participation in previous surveys. The only variable with opposed effects is age, which impact varies from positive to negative in different studies. This could be related to the samples varying in their concentration in different ages, as well as their different locations: the two studies in the Netherlands found positive effects, while the one in Spain found a negative impact. Further, the studies in the Neth-

[3] In this study, the authors focused on willingness and participation. All of those stating willingness participated in sharing photos.

erlands used probability-based panels, while the study in Spain used an opt-in panel. Indeed, more research is needed to understand such differences.

Research Questions and Hypotheses

This paper reports the results of an experiment conducted among online opt-in panel respondents of a mobile survey gathering data on the books they have at their main residence through conventional or image-based response formats. The image-based format involved requesting photos of the books, while the conventional format asked to answer questions related to the books by typing numbers or clicking a radio button. I compare three groups: 1) Conventional format: respondents are asked to answer 11 questions related to the books they have at home by typing in numbers of clicking radio buttons; 2) Image-based format: respondents are requested to send photos of the books at their home; 3) Choice: respondents can decide either to answer the 11 conventional questions or to share photos.

While most of previous research focused on relatively straightforward tasks, involving a single answer or photo, this paper explores the implications of employing conventional versus image-based response formats when interested in more complex tasks. On the one hand, the conventional task is much more demanding than what has been tried in previous studies comparing both response formats. Indeed, it requires participants to answer 11 cognitively demanding questions, that require estimating numbers (e.g., number of books of different categories) and percentages (e.g., proportion of books in Spanish). On the other hand, the image-based response format is also more demanding than in previous research: instead of requesting a single photo (e.g., of the heating system as in Ilic et al. 2022, or of the surroundings as in Bosch et al. 2019), participants were asked to provide photos of all the books in their residence. Thus, respondents might need several photos to capture all books. Further, respondents might need to move through different spaces/rooms within their household, since books might be dispersed (e.g., children have books in their rooms, or books that are being currently read are on night tables). If other people are using some of the rooms (e.g., children sleeping in their room), it might also not be possible to take the photos immediately. Finally, respondents were instructed to remove items such as decorative elements, to ensure clear visibility of the books. This can represent quite some work for participants to remove and put back such items, and can generate high burden if they have books in several places and have to do it several times. Consequently, this study presented respondents with a challenging task, expected to be more time-consuming and effort-intensive than previous examples studied in image-based data collection. Given these complexities, this research could provide novel insights into aspects

previously explored such as the respondents' preferences, participation, and evaluation, but in a slightly different context.

This study first delves into the respondents' preferences concerning response formats when given the choice between conventional and image-based response ones, especially for addressing demanding tasks. Thus, the first research question is:

RQ1: *Do respondents prefer to provide the information about the books in their dwelling through images or answering questions in conventional ways?*

The only study on respondents' preferences between conventional and image-based response formats (Ilic et al., 2022) found that participants mostly choose images. However, due to the specificity of the task studied in this paper, I expect respondents to be more reluctant to capture and share photos since the books could be in many different places, respondents might need to tidy each area before capturing the photos, may not be at home when answering the survey (thus unable to capture the photos), or could have privacy concerns (e.g., perceive the task as more intrusive than in the case of photos of the heating system). Thus, my first hypothesis posits that, when respondents are in a position of choosing a method, the preference for the conventional response format will surpass that of the image-based one (*H1*).

Second, this study assesses the levels of participation, i.e., participants actually answering the 11 conventional questions (for the conventional format) or sending at least one photo displaying their books (for the image-based format). Further, it investigates whether being able to choose one response format affects participation. Thus, the second research question is:

RQ2: *Does the participation vary between a) image-based versus conventional answer formats, and b) respondents choosing their preferred format versus respondents being only proposed one format?*

Based on previous literature, lower participation is expected from respondents (with or without a choice) for image-based inquiries compared to conventional answer formats (*H2a*). Moreover, participants with a choice are expected to participate more than those without a choice in the case of the image-based response format (*H2b*), while participation levels in conventional formats is expected to be similar across groups (*H2c*), as in Ilic et al. (2022).

Respondents participating in the image-based response format might still dislike it or find it difficult, which might potentially affect the participation in future surveys. Thus, the next research question:

RQ3: *How does the evaluation of respondents about the book-related questions vary between a) image-based versus conventional response formats, and b) respondents choosing their preferred format versus being only proposed one?*

Considering the cognitive effort required to provide accurate answers in the conventional format, which involves tasks like estimating numbers and proportions, in contrast to the familiarity of capturing photos with smartphones or tablets, a task for which most respondents in the same panel declared having the skills (see Iglesias & Revilla, 2024), I anticipate that respondents utilizing the image-based response format will perceive the task as easier than those employing the conventional one (*H3a*). However, respondents may like the conventional format more than the image-based one (*H3b*), given the respondents' familiarity with conventional questions, and that capturing photos, although not difficult, might present practical challenges in the case of the books at home, potentially leading to a more time-consuming and tiresome experience. Moreover, I expect that participants choosing their response format present better evaluations than those unable to choose, regarding both the perception of easiness (*H3c*) and the extent to which they like the tasks (*H3d*).

Finally, certain respondents' characteristics could influence their participation and evaluation of image-based response formats.[4] Thus, my last research question is:

RQ4: *How do the respondents' sociodemographic characteristics, experience as panelist, comfort with new technologies, trust in the confidentially of the answers, and use of mobile devices influence their participation and evaluation in image-based versus conventional response formats?*

To the best of my knowledge, no research has studied the influence of these factors on the evaluation of an image-based format. Therefore, I do not formulate hypotheses in this case but follow an exploratory approach. In contrast, I propose the hypotheses bellow regarding the impact of the different factors on participation, since there is some research on this aspect.

Since the survey implemented in this study did not target old population[5] (most respondents were under 50 years old), age is not expected to significantly impact participation in image-based response formats (*H4a*). Similarly, no effect is expected for gender (*H4b*) and education (*H4c*), in line with previous literature. Conversely, familiarity with the camera included in the mobile device (*H4d*) and sharing photos (*H4e*) are expected to have a positive impact in participation in the image-based format. Although there are mixed findings in the previous literature regarding this aspect, I anticipate that individuals accustomed to using cameras in smartphones and sharing photos will be more inclined to participate: since those using smartphones more often are more familiar with them,

[4] This study aimed to investigate how these factors influenced respondents' preferences for one format over the other. However, such analysis could not be conducted due to the low number of respondents opting for the image-based response format (n=12).

[5] According to data from the Economically Active Population Survey of the Statistics Office of Spain, 99% of children attending primary school have parents of maximum 54 years old. See the section "Data collection" for more details.

and those frequently sharing photos might have fewer privacy concerns, these two aspects could increase participation. Similarly, although not studied in the previous literature, a higher level of comfort with new technologies is expected to boost participation in image-based response formats (*H4f*). Moreover, trust in the confidentiality of the answers is expected to impact positively the sharing of photos (*H4g*). Further, it is expected that households with more children will present lower participation (*H4h*), as it might translate into having more books and eventually in more places, making the task of capturing photos more tiresome. Finally, previous experience as a panelist is expected to negatively impact the participation in image-based requests, as respondents might be more accustomed to conventional formats than innovative ones (*H4i*). Table 8, available in the conclusions, summarizes the hypotheses.

Addressing these research questions, this paper contributes to the existing literature by presenting results on the request for images within the context of a mobile survey, exploring a relevant topic in social sciences and focusing on more complex questions and tasks than what has been studied previously. Further, this study focuses on a specific demographic group, namely parents living with children who attend the first, third, or fifth year of primary school in Spain. This introduces practical challenges, such as limited response time due to parental duties, difficulties in capturing photos amid childcare responsibilities, and potentially less organized living spaces, especially concerning children's books. Moreover, this is the first study collecting images with the ***WebdataVisual*** tool (Revilla et al., 2022), which was developed with the goal of having a more user friendly tool. Finally, the relation with smartphones moves forward swiftly, and technology is more accessible each day to smartphone users. Thus, this study complements the previous literature by contributing a contemporary perspective, recognizing the changing landscape in smartphone usage.

Data and Methodology

To address the research questions, an experimental design was implemented. This experiment is part of a bigger study. In this section, only the relevant elements for this paper will be presented. For a depth review on the overall study design, readers can consult the full study protocol (Iglesias et al., 2023).

Experimental Design and Groups

The experiment aimed to collect information about the books present in respondents' main residences using conventional and image-based response formats.[6]

For the conventional response format, 11 questions grouped in the following three dimensions were asked:

- *Number of books*: four open-ended questions about 1) the total number of books at home, and the number of books 2) for toddlers and children who do not know how to read, 3) for literate children and teenagers, and 4) aimed at a general audience.
- *Language*: three open-ended questions asking for the percentage of books 1) in Spanish, 2) in one of the three co-official languages in Spain (Catalan, Galician, and Euskera), and 3) in other languages.
- *Storage*: four radio-button questions asking whether books are stored 1) on shelves, 2) inside closets or drawers, 3) on center, coffee, or night tables or over a desk, and 4) in other places.

For the image-based response format, respondents were only asked to provide photos of their books, under the assumption that the aforementioned information could be extracted through image classification, i.e., the process of extracting and labeling the information contained in an image (Bandyopadhyay, 2021). Both conventional and image-based questions regarding the 11 items will be referred to as "test questions" in this paper.

Three experimental groups are considered: *Text*,[7] *Images*, and *Choice*. For the sake of simplicity, the names of the two first groups reflect their respective assigned answering format. In the third group (*Choice*), participants could select between the conventional or image-based formats. Throughout this paper, respondents choosing the conventional format are referred to as members of *TextChoice*, while those preferring images are named *ImagesChoice*. Respondents stating no preference were assigned to the image-based format (thus, are considered members of *ImagesChoice*). Table 1 presents a summary of the groups and response formats compared in this paper.

---

[6] This collection will help answering substantive questions regarding children's academic performance in relation to the number of books. Since previous literature has found no impact of e-Books in children's academic performance (Heppt et al., 2022), information on them was not collected.

[7] The design of the full experiment considers two different methods within the conventional response formats: *Text* and *TextPlus*. The only difference between both methods is that in the latter an illustration was provided to respondents to help them estimate the number of books. This is used in a different paper to study whether such illustration can help improve the quality of the answers in conventional formats. Since this does not affect the response format, for the analytical purposes of this paper, respondents in *TextPlus* as well as those in *Text* are all included in the *Text* group.

Table 1 Groups and Response Formats

Group	Response format for the test questions
Text *TextChoice*	11 conventional questions.
Images *ImagesChoice*	Capturing and sending photos of the books at home.

Questionnaire

The questionnaire consisted of up to 65 questions, extending beyond the test inquiries and covering topics such as the sociodemographics of respondents, the characterization of (one of) their child in primary school, activities related to literature engagement, usage of camera-related functions on their mobile devices, comfort with new technologies, and self-assessment of their spatial, mathematical, and verbal abilities. Further, respondents were asked to evaluate their experience when answering conventional or image-based response formats, and to provide additional information such as whether they had technical problems while uploading their photos. For more details, the full questionnaire (in Spanish and in English) is available in the supplementary material 2.

Since photos of the books at home could have been potentially asked to any respondents, a message at the beginning of the questionnaire requested them to answer from home. However, this could not be verified, as respondents did not share geolocation data. Thus, respondents could continue with the survey even if they were not at home.

Moreover, respondents had to complete the survey on smartphones or tablets. This restriction was imposed because taking photos of all books with computers (even laptops) was deemed too inconvenient. Further, the ***WebdataVisual*** tool used to collect the photos only allows capturing them during the survey when using mobile devices. Respondents entering the survey via computers were asked to switch to a smartphone or tablet and were unable to continue if they did not do so.

Sample and Data Collection

The target population consisted of parents of children enrolled in the first, third, or fifth year of primary education in Spain at the moment of the survey. These specific years were selected because changes in Spain's evaluation system (shifting from quantitative to qualitative evaluation) were implemented in those courses at the moment of the survey. Thus, including the second, fourth, and sixth years in the same survey might have impacted the substantive objectives

to be fulfilled with the collected data, as the questionnaire asked for grades in Spanish and mathematics (for details, see the study protocol by Iglesias et al., 2023). Quotas for age, gender and educational level of respondents were used to get a sample similar on these variables to adults with children between 6 to 12 years (the average ages of children attending primary school in Spain). These estimates were derived from the Economically Active Population Survey of the Statistics Office of Spain.[8]

Data were collected in June 2023 through the Netquest opt-in panel in Spain (www.netquest.com), which invites panelists to participate in surveys, and rewards them with points determined by the questionnaire's length (for more information about the kind of surveys and rewards in this panel, see Revilla, 2017).

Out of 4,854 individuals invited to participate, 2,443 started the survey. 899 were filtered out due to security checks or survey requirements not being met (e.g., not providing consent to participate or not having a child in the first, third, or fifth year of primary education), while 72 individuals were excluded because demographic quotas had already been fulfilled. 202 entered to the survey but broke off before the first test question (i.e., first question about the books at home), leading to 1,270 individuals arriving to the test questions: 53% were female, the mean age was 42 years and 92% of participants were 30 to 50 years old. 45% possessed a higher education degree. Of all respondents arriving to the test questions, 636 were in the *Text* group, 305 in the *Choice* group (261 in *TextChoice* and 44 in *ImagesChoice*), and 329 in the *Images* group.

The allocation in a given group was performed right before the first test question, with respondents being assigned to the group with the least individuals at that moment. The group *Text* is two times larger than the others since it contains two groups (see footnote 7). Checks for balance were conducted (see supplementary material 3) on age, gender, and level of education, revealing no differences between the composition of the groups *ImagesChoice* and *Images*, and between *TextChoice* and *Text*. When comparing participants answering through either conventional or image-based format, differences are found for gender, with a significantly higher proportion of women in the image-based than in the conventional group (50% versus 59%). However, the difference between groups

[8] The public dataset (available at https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=resultados&idp=1254735976595#tabs-1254736030639) displays the age of the members of the dwellings in ranges of 5 years. For the estimation aimed to calculate the quotas, dwellings with children between 5-9 years old during the first trimester of 2022 were considered, since those children will be 6-10 years old in the same trimester of 2023. Thus, our quotas are a proxy for dwellings with children between 6 and 12 years old, since it is not possible to know that exact range of ages from the publicly available data. Margins of +/-3 percentage points were used for the quotas, since the target population of this study is not exactly the same as the one used in the Economically Active Population Survey.

being of 9 percentage points, I do not expect it to influence the overall results presented in this paper.

On average, respondents reaching the test questions have been members of the Netquest panel for 6.9 years, and completed 13 surveys in the three months previous to this study. 99% of respondents used a smartphone to answer the survey (1% used a tablet). For those finishing the survey, the median completion time was 9.3 minutes (9.5 for the image-based response format groups, and 9.2 for those answering through conventional questions).

Analyses

R 4.2.3 was used to perform the analyses. The script is available in the project's repository (https://osf.io/7y3sq/).

Addressing ***RQ1*** (preferences for one response-format over the other), respondents in the group *Choice* were asked twice for their preferred response format and were offered three options: conventional, image-based, or no preference. The first question regarding preference was presented before the test questions so participants could answer those questions by using the response format they chose (those without a preference were assigned to answering with photos). The second preference question was presented after they answered the test questions with the chosen format, to assess whether they would still prefer it. The proportions of participants selecting each option within those who saw the questions are reported. Comparisons are made between the three categories to test *H1*.

Further, respondents not choosing images were asked for their reasons through a multiple-choice question with the following options: camera in the mobile device not working, privacy concerns, expected lack of skills, having too many books, and others (with the option of explaining further in a textbox). The main reasons are presented.

To study participation (***RQ2***), different dimensions were considered. Indeed, when facing a given question, participants have three main options: provide an answer (participation), skip the question but continue with the survey (item nonresponse) or abandon the survey (breakoff). In this study, the interest is in comparing a set of 11 questions with a request for photos. Thus, there are different ways to operationalize breakoff, item nonresponse and participation, in each response format. Consequently, I computed and report several indicators, which were estimated separately for respondents answering conventional and image-based formats.

As for the conventional one, five indicators are presented:

- *Breakoff*: Percentage of respondents, within those who saw the first question about books, that left the survey on one of the screens where the 11 test questions were displayed.

- *Minimum participation*: Percentage of respondents providing a substantive answer to at least one question.
- *Partial participation:* Percentage of respondents providing a substantive answer to at least six questions.
- *Full participation:* Percentage of respondents providing a substantive answer to all 11 questions.
- *Average number of substantive answers* out of the 11 possible ones.

In the last four indicators the calculations are computed out of all respondents seeing the 11 questions and continuing with the survey (i.e., not breaking off). Further, the option "I don't know" was presented to participants answering through conventional formats. Even though this might be a valid response, particularly when participants genuinely lacked the information, I excluded "I don't know" when studying participation, because this can be selected as a way to avoid any cognitive effort, and because there is no equivalent for the image-based response format. Conclusions reached in this paper do not change when considering "I don't know" as participation (see supplementary material 4).

For the three questions concerning the language(s) of the books (the percentages of books in Spanish, in one of the co-official languages in Spain, and in other languages), if the answered questions added up to 100, the three items were considered as complete. For example, if a respondent had all their books in Spanish, they might have typed "100" for Spanish and left the others blank. This is considered as a participation without item nonresponse. In any other case where there was a blank response without adding up to 100, the empty questions were considered as nonresponse.

As for the image-based format, three indicators were used:

- *Breakoff:* Percentage of respondents leaving the survey on the screen where the image request was presented over the number of respondents who saw this screen.
- *Minimum participation:* Percentage of respondents capturing and sharing at least one image. Since it is not possible to know if one photo captured all the books in the dwelling, sending at least one image was considered as "minimum participation". The denominator was the number of respondents required for images who did not breakoff in the test question.
- *Average number of photos* among participants sharing at least one image.

Comparisons were made at both the response format (conventional versus image-based) and group (*Text* versus *TextChoice, Images* versus *ImagesChoice*) levels. Regarding response format, the percentages of respondents breaking off were compared to test *H2a*. The rest of indicators used in both formats cannot be directly compared since these measures gauged different aspects. For instance,

answering one test question in the conventional response format (minimum participation) is not equivalent to sending one photo.

At the group level, I compared the *Images* and *ImagesChoice* groups based on their percentages of breakoff, respondents sending at least one photo, and the average number of photos to test *H2b*. Similarly, the *Text* and *TextChoice* groups were compared regarding their percentages of breakoff, respondents providing at least one, six and 11 substantive answers, and the average number of such answers to test *H2c*.

Further, the reasons for not uploading images, asked to participants skipping the question requesting images, are reported. The same categories presented to those not choosing images in the preference question were offered, with an additional category for technical issues.

Regarding ***RQ3***, two aspects of the respondents' evaluation were considered: the extent to which they found the test questions easy/difficult, and how much they liked/disliked them. These aspects were originally measured through a scale from 0 ("Extremely difficult"/ "Totally disliked") to 4 ("Extremely easy"/ "Totally liked"), and were recategorized into "Difficult"/ "Dislike", "Not easy nor difficult"/ "Not like nor dislike", and "Easy"/ "Like". The proportions of respondents in each of these categories over those presented with these questions are compared among response formats and groups to test *H3a* to *H3d*.

Further, an open-narrative question among those disliking any of the two response formats was presented. The answers to these questions were coded and the frequency of the codes was estimated (n=20 for images, n=15 for conventional questions). Only codes mentioned more than once are presented. With such small groups conclusions cannot be reached, but the reasons still help understanding why respondents did not like the respective response formats.

Comparisons between categories of a variable within the same group (***RQ1***) and comparisons among groups and formats (***RQ2*** and ***RQ3***) were performed with Chi-squared tests, with significance at the 5% level.

Finally, regarding ***RQ4***, logistic regression analyses were performed to assess the extent to which participation and evaluation of the test questions were impacted by the respondents' characteristics. These characteristics included gender (1=female, 0=male), age (numerical), level of education (1=tertiary education, 0=secondary education or less), number of children (numerical), frequencies of camera use and images sharing (1=at least once a week, 0=less often), experience as a Netquest panelist (logarithm of the number of surveys completed in the last three months), comfort with new technologies (1=very or totally comfortable, 0=not at all to quite comfortable) and trust in the confidentiality of the answers (1=trust, 0=no trust).

For evaluation, liking the survey and finding it easy were used as dependent variables. Regarding participation, the dependent variable for the conventional format was the full participation, and for images the minimum participation.

The former was chosen as it presented the ideal scenario in the conventional format, i.e., answering all the 11 questions. The minimum participation was selected for the image-based format because even one photo has the potential to contain all the information of interest, making it the minimum standard for image submissions. These two regressions were employed to test *H4a* to *H4i*.

Results

Respondents' Preferences

To address ***RQ1***, Table 2 presents the preferences of respondents in the *Choice* group before the test questions and after.

Table 2 Preferences of Respondents in Choice Group Before and After Test Questions (in %)

	Preference for…		
Group	Conventional (a)	Image-based (b)	No preference (c)
Before seeing the test questions			
Choice (n=305)	85.6[b,c]	3.9[c]	10.5
After seeing the test questions and having stated a preference			
Initially preferred conventional (n=258)	91.1[b,c]	1.6[c]	7.4
Initially preferred image-based (n=12)	33.3	50.0	16.7
No initial preference (n=31)	16.1[c]	12.9[c]	71.0

Note: Letters in superscript specify the statistically significant differences between categories.

Among respondents who had the option to choose a response format before the test questions, a clear preference emerged: 85.6% favored the conventional format, while only 3.9% preferred the image-based one. These results support *H1*. Another 10.5% expressed no particular preference, which resulted in the *Images-Choice* group being composed of more respondents without a preference than of respondents actively choosing images.

When asking for the reasons for not choosing images to those who preferred conventional questions (261 cases), respondents mainly stated having an extensive book collection and being reluctant to photograph all of them (43.8%), and concerns related to privacy (39.2%).

Regarding preferences after seeing the test questions, 91.1% of participants using the conventional format would choose it again, while only half of those

using images would do so. Still, the number of respondents choosing images was very small (n=12), which prevents reaching conclusions on this matter. Finally, 71.0% of those without a preference, who were assigned to answering with images, still did not state a preference after answering the test questions, and 12.9% would choose images after having used them to answer.

Participation

Concerning the participation of respondents in conventional and image-based response formats (***RQ2***), Table 3 presents the breakoff rates, while Table 4 displays the indicators of full, partial, and minimum participation, and the average number of answers responded and photos sent by participants.

Table 3 Breakoff Rate by Response Format and Group (%)

By…	Sample size	Breakoff rate
Response format		
Conventional	897	**0.6**
Image-based	373	7.5
Group		
Text	636	0.6
TextChoice	261	0.4
Images	329	8.2
ImagesChoice	44	2.3

Note: Bold notes statistically significant differences among formats or groups.

The percentage of breakoff is significantly lower among respondents using the conventional format (0.6%) compared to those asked to send images (7.5%). When comparing the groups, there are no significant differences between those with and without a choice, and there is an inclination for breakoff to be more distinct among the images groups: 8.2% in the *Images* group broke off, while 2.3% in the *ImagesChoice* group did so, but the difference is not statistically significant.

Table 4 Other Indicators of Participation by Response Format and Group

By…	Sample size	Type of participation (%)			Average number of answers/photos
		Minimum	Partial	Full	
Response format					
Conventional	892	100	99.9	79.5	10.5
Image-based	345	39.7			2.9
Group					
Text	632	100	99.8	78.5	10.4
TextChoice	260	100	100	81.9	10.6
Images	302	**37.7**			2.9
ImagesChoice	43	**53.5**			2.7

Note: Bold notes statistically significant differences among formats or groups answering through the same format.

As per the different levels of participation in the conventional format, almost all respondents answered at least half of the questions (partial participation), with no major differences among the two conventional groups, and 79.5% of respondents answered the 11 questions (full participation). Although there are no significant differences among groups, the *TextChoice* group has a slightly higher proportion of respondents providing all answers (3.4 pp). Finally, the average number of answered questions is 10.5, with no significant differences among groups.

Regarding the image-based format, 39.7% sent at least one image (minimum participation). In this case, there are statistically significant differences among groups, with 53.5% of those in *ImagesChoice* providing images, against 37.7% among those without a choice in the group *Images*. Finally, the average number of photos per respondent among those actually sending photos (i.e., excluding the 60.3% who did not send the photos when required) is 2.9 photos, ranging from 1 to 16 photos per respondent, and without statistically significant differences between the two image-based groups. Considering all respondents asked for photos (also those who did not send any), the mean number of images drops to 1.2. In all cases, these photos might or might not cover all the books at the residence.

As for the reasons for not sending photos, respondents mentioned privacy concerns (43.0%), technical issues when uploading the photos (13.5%), camera in the device not working (10.6%), and lack of skills (10.1%). In the open-ended question, 11.1% of respondents explained that they were not at home.

Overall, these results confirm *H2a* (lower participation in the image-based format compared to the conventional one). Furthermore, *H2b* and *H2c* are also supported, as significantly higher proportions of participants provided images

in the *ImagesChoice* than in the *Images* group, whereas results were more similar between *Text* and *TextChoice*.

Evaluation of the Test Questions

To address ***RQ3,*** respondents' evaluations are presented in Table 5.

Table 5 Easiness/Difficulty and Like/Dislike by Response Format and Group (in %)

	Response formats		Groups			
Categories	Conventional (n=891)	Image-based (n=135)	*Text* (n=632)	*TextChoice* (n=259)	*Images* (n=112)	*ImagesChoice* (n=23)
Easy	64.4	66.7	**61.4**	**71.8**	63.4	82.6
Not easy nor difficult	27.2	26.7	29.0	22.8	28.6	17.4
Difficult	8.4	6.7	**9.7**	**5.4**	8.0	0
Like	**53.5**	**24.4**	52.2	56.8	21.4	39.1
Not like nor dislike	**44.8**	**60.7**	45.9	42.1	61.6	56.5
Dislike	**1.7**	**14.8**	1.9	1.2	17.0	4.3

Note: Bold notes statistically significant differences among formats or groups answering through the same format.

For the easiness/difficulty to answer the test questions, no significant differences are observed between response formats, with most respondents finding it easy to answer the 11 questions (64.4%) as well as capture and send photos (66.7%). These results do not support *H3a*.

Regarding groups, respondents in both *TextChoice* and *ImagesChoice* tend to perceive both formats as easier compared to participants in the equivalent non-choice groups, indicating that offering the option to choose leads to a more positive perception on the ease of both response formats. However, results are significant only for the groups using the conventional format, providing partial support for hypothesis *H3c*. Additionally, the group with the highest prevalence in the category "Easy" is *ImagesChoice*. Further, the perception of easiness of participants in groups without a choice is very similar (63.4% for *Images* and 61.4% for *Text*).

Stronger differences are found when examining the extent to which respondents (dis)liked answering these questions. 53.5% of the respondents liked answering the conventional questions and 1.7% disliked it. In contrast, only 24.4% of those answering through images liked it and 14.8% expressed dislike. Further, there were high levels of indifference (60.7% of "not like nor dislike") among the image-based response format. All the categories present statistically

significant differences in favor of conventional questions, providing support for *H3b*.

As per the groups, those who could choose presented higher levels of liking compared to those without a choice, although the differences are not statistically significant. The variation between groups was particularly pronounced for those answering through images: liking among *ImagesChoice* respondents (39.1%) was 18 percentage points higher than in the *Images* group (21.4%). The lack of statistically significant findings does not support H3d.

Of those disliking capturing and sending photos (n=20), 12 individuals reported privacy concerns. As for the conventional format, five of the 15 respondents expressing dislike mentioned they chose "dislike" by mistake, and three found it too difficult, burdensome or time consuming to answer all the questions regarding books. Thus, the reasons for not liking both formats vary.

Impact of Respondents' Characteristics

Impact on Participation

Regarding the impact of participants' characteristics (***RQ4***), Table 6 presents the results of logistic regressions explaining the full participation for the conventional response format (i.e., answering the 11 questions) and minimum participation for the image-based format (i.e., sending at least one photo).

Table 6 Logistic Regressions for Participation

	Participation (=1)	
	Conventional	Image-based
Female	-0.157	0.653**
Age	0.008	0.046**
Tertiary education	0.500***	0.530**
Number children	-0.067	-0.213
Using camera at least once a week	0.559**	0.578*
Sharing images at least once a week	-0.188	-0.045
Number surveys last three months	0.455	0.109
Comfortable with new technologies	0.418**	0.380
Trust confidentiality	0.400**	0.368
Constant	-0.070	-3.438***
n	854	341
Log Likelihood	-412.151	-213.047

Note: *p<0.1; **p<0.05; ***p<0.01

First, having a higher education degree, and using the camera at least once a week have significant and positive effects on participation in both conventional and image-based response formats. These results contradict *H4c* (stating no influence of education) but support *H4d*.

Second, other variables have significant effects for only one response format. For example, while identifying as female affects positively and significantly the participation in image-based response formats, the effect is not significant in the conventional one. Therefore, female participants were more inclined to capture and share images, but gender did not play a role when it came to answering the conventional questions. Since gender influenced at least one of the formats, hypothesis *H4b* is not supported. Similarly, age is a significant factor only for the image-based response format: an older age is associated with higher participation in the questions asking for photos, contradicting *H4a*.

Higher levels of comfort with new technologies and trust in the confidentiality of the answers significantly impact the participation in the conventional response format, but not in the image-based format, not supporting *H4f* and *H4g*.

Finally, the frequency of sharing photos, the number of children, and the number of Netquest surveys answered before this study do not influence participation in either format, thus *H4e, H4h,* and *H4i* are not supported.

Impact on Evaluation

As per the evaluation of the test questions, Table 7 shows the results of the logistic regressions for finding them easy and liking them.

First, education, number of children and frequency of sharing images did not significantly influence the perception of easiness and liking, neither in conventional nor image-based response formats.

Second, some variables influence the perception of easiness or liking but just of one response format. Using the camera at least once a week negatively impacted the liking of the conventional format. Further, being female only influenced (in a positive way) the perception of easiness of the image-based format.

Other variables, as age and trust in confidentiality, affected both dimensions of only one format. While being older decreased the likelihood of liking and finding the questions in the conventional format easy, trusting in the confidentiality of the answers made it more likely.

Finally, some variables impacted both response formats. Feeling comfortable with new technologies increased the likelihood of liking the two formats, and finding the conventional format easy. Further, the number of Netquest surveys completed in the three months prior to this study had significant positive effects on the respondents' perceived ease of the test questions, and in the likability of the conventional format.

Table 7 Logistic Regressions for Easy and Like

	Easy (=1)		Like (=1)	
	Conventional	Image-based	Conventional	Image-based
Female	-0.189	0.742*	-0.003	0.617
Age	-0.026*	-0.004	-0.041***	-0.067
Tertiary education	-0.100	-0.407	0.087	-0.071
Number children	0.004	-0.219	-0.078	0.395
Using camera at least once a week	0.031	0.614	-0.460**	-0.171
Sharing images at least once a week	-0.172	0.235	0.138	0.793
Number surveys last three months	0.644***	1.698**	0.723***	1.037
Comfortable with new technologies	0.636***	0.557	0.444***	1.064**
Trust confidentiality	0.414***	0.393	0.456***	0.722
Constant	0.764	-1.776	1.071	-2.103
n	854	135	854	135
Log Likelihood	-530.533	-77.286	-563.897	-61.894

Note: *p<0.1; **p<0.05; ***p<0.01

Conclusions

Summary of Main Results

In this paper, the focus was on the respondents' preferences, participation, and evaluation of questions answered though conventional or image-based format, and the impact of respondents' characteristics on their participation and evaluation of both formats. Table 8 presents a summary of the hypotheses and their support based on the findings of this study.

First, a clear preference among respondents for the conventional format over the image-based one was found (***RQ1***). These results were very conclusive as only 4% of participants opted for sending photos. These findings contradict the only study investigating respondents' preferences, by Ilic et al., 2022, where most respondents opted for images. This could be due to the task in this paper being more demanding than the one conducted by Ilic et al., (2022), who asked for one photo of three places within the household.

Table 8 Summary of Hypotheses and Their Support

Hypotheses	Result
H1: Higher preference for conventional response format.	Supported.
H2a: Lower participation in image-based format.	Supported.
H2b: Higher participation in the image-based response format when possible to choose.	Supported.
H2c: Participation in conventional format not affected by having a choice.	Supported.
*H3*a: Respondents using the image-based format perceive the test questions as easier.	Not supported
H3b: Respondents like the conventional format better than the image-based one.	Supported.
H3c: Participants with a choice find the test questions easier than those without a choice.	Partly supported (statistically significant results only for the conventional format groups).
H3d: Participants with a choice like the test questions better than those without a choice.	Not supported
No effect on participation in image-based response format of: *H4a*: age, *H4b*: gender, *H4c*: education.	Not supported (positive effect on participation for the image-based format).
Positive effect on participation in image-based format of: *H4d*: familiarity with the device camera, *H4e*: sharing photos with the device, *H4f*: being comfortable with new technologies, *H4g*: trusting the confidentiality of the answers.	*H4d* supported. *H4e*, *H4f* and *H4g* not supported
Negative effect on participation in image-based response format of: *H4h*: more children in the household, *H4i*: higher participation in previous surveys.	Not supported

Second, participation (***RQ2***) was lower among image-based format respondents with only 40% sending photos, compared to 80% of participants in the conventional format answering all questions. Participation in questions asking for images was lower than in some previous studies (55% in Bosch et al., 2019, 49-67% in Bosch et al., 2022), but in line with what was found by Ilic et al. (2022),

where less than 40% of respondents in three out of six groups complied with the task. Additionally, participation was below the expected participation when it comes to sending photos of something in the house (54% in Iglesias & Revilla, 2024, also using the Netquest panel in Spain). The variance between the actual and expected participation rates might stem from the different target populations and the increased complexity of the tasks assigned in this study.

Moreover, the option to choose significantly influenced participation rates in questions requesting photos, with 54% of respondents in the *ImagesChoice* group sending at least one (compared to 38% among those in the group *Images,* without a choice), although the sample size of the group abstaining from conventional questions was small. The participation rates were similar to the study by Ilic et al. (2022) (between 50%-78% for those choosing images, and 27% and 39% for those automatically assigned to send photos). Being able to choose did not result in significant differences in participation for conventional questions, also similar to Ilic et al. (2022). The reasons for refraining from sharing photos were predominantly linked to privacy concerns, technical challenges, and participants not being at home during the survey.

Third, the perception of easiness of the test questions was similar between the two formats (***RQ3***), but the conventional format was liked better. Further, respondents in *TextChoice* found the questions about books easier compared to those in the *Text* group. *ImagesChoice* respondents also found the question easier than those without a choice in *Images*, but these results were not significant.

Finally, concerning the factors influencing the participation (***RQ4***), age, being female, and using the camera in the mobile device at least once a week increased the likelihood of participating in the image-based response format questions. Unlike previous literature (Iglesias & Revilla, 2024; Struminskaya, Lugtig, et al., 2021), counting with tertiary education also had a positive effect on participation in the image-based format. However, the number of children in the household, frequency of sharing images, number of Netquest surveys completed, comfort with new technologies, and trust in the confidentiality of the answers did not demonstrate any significant impact.

Regarding the evaluation, being female and completing more Netquest surveys made it more likely to find the image-based format easy, while feeling comfortable with technology favored the liking of this type of questions. Age, number of children in the household, education, trust in the confidentiality of the answers, and frequency of capturing and sharing images did not impact the evaluation of image-based questions.

Limitations

The study has some limitations. First, data were collected in an opt-in panel in Spain. Different results could be obtained in other types of panels or places. Fur-

ther, respondents in this panel are used to answering questions in conventional formats, which could partly explain their higher participation in that format rather than in the image-based one.

Second, the targeted population was very specific: parents of children attending the first, third, or fifth year of primary school. Having such a specific population might create additional challenges (e.g., respondents less able to capture photos since they are taking care of their children), thus findings from this paper should not be generalized to other topics (i.e., photos of things other than books at home) or to other populations without carefully considering the similarities and differences with the target population and topic of this study. Still, researchers can use these results as a starting point to plan the collection of photos in other settings.

Third, the quality of the information collected through the images and the conventional questions was not assessed in this study. Respondents in the conventional format might have provided approximate answers, not have considered the books of all the members, or invented answers to finish the survey more quickly. Participants sending photos might have photographed only part of the books or sent off-topic photos. Thus, analyses of data quality are needed.

Finally, a significant number of respondents completed the survey from locations other than their home, even if a message asking to answer from home was presented at the beginning of the survey. This predominantly impacted participants using the image-based format, as they could not capture and send real-time photos while answering the survey. Similarly, conventional format respondents willing to count or refer to the books to provide a more accurate answer could not do so.

Practical Implications

The outcomes of this study provide valuable insights for guiding future research endeavors involving photo collection through online surveys. First, participation in questions asking for images is likely to be lower compared to conventional questions. Thus, researchers should balance whether the content of the images obtained outweighs the potentially lower participation.

Second, if researchers are interested in continuing with photo collection, they should consider strategies to improve participation. One approach could be mentioning the reward for sharing photos to participants when presenting the question, allowing them to assess the potential benefits of capturing and sending images. Further, when possible, participants might be informed before the survey about the photos they will be asked for, enabling them to answer the questionnaire when they are able to capture such photos.

Third, although a small part of respondents preferred sending images over the conventional format, having a choice made a difference in terms of partici-

pation. This could lead researchers to present respondents with the opportunity to choose to increase participation, but the implications of requiring images (e.g., longer implementation or fees associated to storing images) should be considered before making such a decision involving a potentially limited part of their sample. Given the non-negligible risk of low participation and the great efforts to implement the collection of photos in surveys, maybe we as researchers are not yet ready to replace conventional questions with photos, but we might combine both formats to have additional/complementary information. In this case, it would mean asking for the characteristics of books in conventional ways and additionally ask for the photos of the books.

Fourth, respondents liked the conventional format better than the image-based one, which reinforces the idea of carefully deciding when it is worth it to ask for images (weighting the benefits and disadvantages). Researchers should consider options to make the overall experience more likable, such as giving the option of capturing the photos whenever the respondent wishes to do so (e.g., letting them know before the survey).

Finally, the high number of responses per participant in the conventional format (10.5 out of 11) poses an optimistic scenario, suggesting that respondents will answer questions, even if they are cognitively demanding. However, such answers might be (deliberately or not) incorrect, given the difficulty associated to answer with accuracy the number of books per category and the percentages of books written in certain languages.

Overall, researchers should carefully consider when and how to ask for images in a survey, balancing the benefits of this format (i.e., potential better quality and types of insights) with its disadvantages (lower participation, investment in resources and time). These factors should also be compared with the expected outcomes of conventional questions, in order to decide which type of questions work better, or even consider using both formats to compensate their drawbacks and promote their benefits: combining the two formats might lead to higher participation rates, since respondents would answer the conventional questions and potentially also send photos, which could allow gaining details regarding the number of books, extracting other information of interest, and also assessing the accuracy and quality of the answers provided in conventional ways.

Data availability statement

The script and supplementary materials are available in the project's repository (https://osf.io/7y3sq/). The dataset will be available in the same repository soon after the main papers related to this project are accepted.

Acknowledgements

I am very grateful to Melanie Revilla for her valuable and constant feedback in the development of this paper, and to her, Birgit Heppt, Anna Volodina, and Clemens Lechner for their collaboration on the design of this research. Also, I would like to thank Carlos Ochoa, Oriol J. Bosch and Agustín González for their review and comments on the questionnaire, and María Paula Acuña and Ksenija Ivanovic for their help in checking the programming of the survey.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 849165), the Agencia Nacional de Investigación y Desarrollo (ANID) under the "Becas Chile" Doctoral Fellowship programme (grant No. 72220301 to Patricia A. Iglesias), and GESIS – Leibniz-Institut für Sozialwissenschaften.

Conflict of interest

The author declared no potential conflicts of interest regarding the research, authorship, and/or publication of this article.

Ethical approval

The WEB DATA OPP project, from which this is study is part, was reviewed and approved by the Institutional Committee for Ethical Review of Projects from the Universitat Pompeu Fabra.

Informed consent

All participants were presented with an information sheet before starting and only those providing informed consent could participate in the survey.

References

Bandyopadhyay, H. (2021, July 6). *Image Classification Explained: An Introduction*. V7. https://www.v7labs.com/blog/image-classification-guide

Bosch, O., Revilla, M., & Paura, E. (2019). Answering mobile surveys with images: An exploration using a computer vision API. *Social Science Computer Review, 37*(5), 669–683. https://doi.org/10.1177/0894439318791515

Bosch, O., Revilla, M., Qureshi, D., & Höhne, J. K. (2022). A new experiment on the use of images to answer web survey questions. *Journal of the Royal Statistical Society, 185*(3), 955–980. https://doi.org/10.1111/rssa.12856

Brunello, G., Weber, G., & Weiss, C. T. (2012). Books are forever: Early life conditions, education and lifetime earnings in Europe. *ISER Discussion Paper, No. 842, Osaka University, Institute of Social and Economic Research (ISER)*. https://www.econstor.eu/bitstream/10419/92617/1/715521357.pdf

Gil Flores, J. (2011). Medición del nivel socioeconómico familiar en el alumnado de Educación Primaria. *Revista de Educación, 362*. https://doi.org/10.4438/1988-592X-RE-2011-362-162

Güre, Ö. B., Şevgi̇n, H., & Kayri̇, M. (2023). Reviewing the Factors Affecting PISA Reading Skills by Using Random Forest and MARS Methods. *International Journal of Contemporary Educational Research, 10*(1), 181–196. https://doi.org/10.33200/ijcer.1192590

Heppt, B., Olczyk, M., & Volodina, A. (2022). Number of books at home as an indicator of socioeconomic status: Examining its extensions and their incremental validity for academic achievement. *Social Psychology of Education, 25*(4), 903–928. https://doi.org/10.1007/s11218-022-09704-8

Höhne, J. K., & Gavras, K. (2022). *Typing or Speaking? Comparing Text and Voice Answers to Open Questions on Sensitive Topics in Smartphone Surveys* (SSRN Scholarly Paper 4239015). https://doi.org/10.2139/ssrn.4239015

Iglesias, P. A., Ochoa, C., & Revilla, M. (2024). A practical guide to (successfully) collect and process images through online surveys. *Social Sciences & Humanities Open, 9*, 100792. https://doi.org/10.1016/j.ssaho.2023.100792

Iglesias, P. A., & Revilla, M. (2024). Skills, availability, willingness, expected participation and burden of sharing visual data within the frame of web surveys. *Quality & Quantity, 58*(2), 1071–1092. https://doi.org/10.1007/s11135-023-01670-3

Iglesias, P. A., Revilla, M., Heppt, B., Volodina, A., & Lechner, C. (2023). Protocol for a web survey experiment studying the feasibility of asking respondents to capture and submit photos of the books they have at home and the resulting data quality. *Open Research Europe, 3*(202), 1–14. https://doi.org/10.12688/openreseurope.16507.1

Ilic, G., Lugtig, P., Schouten, B., Streefkerk, M., Mulder, J., Kumar, P., & Höcük, S. (2022). Pictures instead of survey questions: An experimental investigation of the feasibility of using pictures in a housing survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 185*(Supplement_2), S437–S460. https://doi.org/10.1111/rssa.12960

McNally, S., Leech, K. A., Corriveau, K. H., & Daly, M. (2023). Indirect Effects of Early Shared Reading and Access to Books on Reading Vocabulary in Middle Childhood. *Scientific Studies of Reading*, 1–18. https://doi.org/10.1080/10888438.2023.2220846

Ochoa, C. (2022). Willingness to participate in geolocation-based research. *PLoS ONE, 17*(12). https://doi.org/10.1371/journal.pone.0278416

Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, *9*(2), 293–313. https://doi.org/10.1177/2050157920959106

Revilla, M. (2017). Analyzing Survey Characteristics, Participation, and Evaluation Across 186 Surveys in an Online Opt-In Panel in Spain. *Methods, Data, Analyses*, *11*(2), 135–162. https://doi.org/10.12758/mda.2017.02

Revilla, M. (2022). How to enhance web survey data using metered, geolocation, visual and voice data? *Survey Research Methods*, *16*(1), 1–12. https://doi.org/10.18148/srm/2022.v16i1.8013

Revilla, M., & Couper, M. P. (2021). Improving the Use of Voice Recording in a Smartphone Survey. *Social Science Computer Review*, *39*(6), 1159–1178. https://doi.org/10.1177/0894439319888708

Revilla, M., Couper, M. P., & Ochoa, C. (2019). Willingness of online panelists to perform additional tasks. *Methods, Data, Analyses*, *13*(2), 223–252. https://doi.org/10.12758/mda.2018.01

Revilla, M., Iglesias, P. A., Ochoa, C., & Antón, D. (2022). *WebdataVisual: A tool to collect visual data within the frame of web surveys* [Computer software]. OSF. https://doi.org/10.17605/OSF.IO/R7CAX

Sanders, L. M., Zacur, G., Haecker, T., & Klass, P. (2004). Number of Children's Books in the Home: An Indicator of Parent Health Literacy. *Ambulatory Pediatrics*, *4*(5), 424–428. https://doi.org/10.1367/A04-003R.1

Saris, W., & Gallhofer, I. (2014). *Design, evaluation, and analysis of questionnaires for survey research* (Second edition). Wiley.

Sewall, C. J. R., Goldstein, T. R., Wright, A. G. C., & Rosen, D. (2022). Does Objectively Measured Social-Media or Smartphone Use Predict Depression, Anxiety, or Social Isolation Among Young Adults? *Clinical Psychological Science*, *10*(5), 997–1014. https://doi.org/10.1177/21677026221078309

Sieben, S., & Lechner, C. M. (2019). Measuring cultural capital through the number of books in the household. *Measurement Instruments for the Social Sciences*, *1*(1), 1–6. https://doi.org/10.1186/s42409-018-0006-0

Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., & Dolmans, R. (2021). Sharing Data Collected with Smartphone Sensors. *Public Opinion Quarterly*, *85*(S1), 423–462. https://doi.org/10.1093/poq/nfab025

Struminskaya, B., Toepoel, V., Lugtig, P., Haan, M., Luiten, A., & Schouten, B. (2021). Understanding willingness to share smartphone-sensor data. *Public Opinion Quarterly*, *84*(3), 725–759. https://doi.org/10.1093/poq/nfaa044

Wenz, A., Jäckle, A., & Couper, M. P. (2019). Willingness to use mobile technologies for data collection in a probability household panel. *Survey Research Methods*, *13*(1), 1–22. https://doi.org/10.18148/srm/2019.v1i1.7298

Appendix

Supplementary Material 1: Variables Affecting the Willingness to Share Visual Data

Table 1 Summary of the Effect of Respondents' Characteristics on Willingness to Share Sensor-Collected Data

Variables	No effect	Positive effect	Negative effect
Age		Struminskaya, Lugtig, et al., 2021 Struminskaya, Toepoel, et al., 2021 (65+ y/o)	Iglesias and Revilla, 2024
Female	Iglesias and Revilla, 2024 Struminskaya, Toepoel, et al., 2021	Struminskaya, Lugtig, et al., 2021	
Education	Iglesias and Revilla, 2024 Struminskaya, Lugtig, et al., 2021 Struminskaya, Toepoel, et al., 2021		
Trust in the anonymity of answers		Struminskaya, Toepoel, et al., 2021	
Frequency of taking photos	Struminskaya, Lugtig, et al., 2021	Struminskaya, Toepoel, et al., 2021	
Participation in previous surveys	Struminskaya, Lugtig, et al., 2021	Iglesias and Revilla, 2024	

Note: Empty boxes mean no findings in the previous literature. The results by Struminskaya, Lugtig, et al. (2021) can be applied to actual participation, since all participants who were willing to participate also shared photos.

Table 2 Summary of the Effect of Respondents' Characteristics on the Expected Participation to Share Sensor-Collected Data

Variables	No effect	Positive effect	Negative effect
Age			Iglesias and Revilla, 2024
Female	Iglesias and Revilla, 2024		
Education	Iglesias and Revilla, 2024		
Trust in the anonymity of answers			
Frequency of taking photos			
Participation in previous surveys		Iglesias and Revilla, 2024	

Note: Empty boxes mean no findings in the previous literature.

Supplementary material 2: Fieldwork Document

Download fieldwork document

Supplementary Material 3: Balance Checks for the Experimental Groups

Table 1 Proportion of Categories in Sociodemographic Variables for Participants Assigned to a Group, per Format (in %)

Category	Conventional format (*n=897*) (a)	Images-based format (*n=373*) (b)
Female (vs. male)	50[b]	59
40 y/o or more (vs. 18-39 y/o)	67	62
Tertiary education (vs. low and middle education)	46	44

Note: letters in superscript specify the statistically significant differences.

Table 2 Proportion of Categories in Sociodemographic Variables for Participants Assigned to a Group, per Group Combining Format and Preference (in %)

Category	*Text* *(n=636)* (a)	*TextChoice* *(n=261)* (b)	*Images* *(n=329)* (c)	*ImagesChoice* *(n=44)* (d)
Female (vs. male)	51	49	59	57
40 y/o or more (vs. 18-39 y/o)	66	70	61	68
Tertiary education (vs. low and middle education)	44	49	44	46

Note: Since the analyses are performed between groups of the same format, statistical comparisons were not conducted among groups of different methods. In this table, no statistically significant differences were observed.

Supplementary Material 4: Participation Indicators Considering "Don't Know" as Participation

Table 1 Breakoff Rate by Response Format and Group, Considering "Don't Know" as Participation

By…	Sample size	Breakoff rate
Response format		
Conventional	897	**0.6**
Image-based	373	7.5
Group		
Text	636	0.6
TextChoice	261	0.4
Images	329	8.2
ImagesChoice	44	2.3

Note: bold notes statistically significant differences among formats. No statically significant differences were found among groups.

Table 2 Other Indicators of Participation by Response Format and Group, Considering "Don't Know" as Participation

By…	Sample size	Type of participation (%)			Average number of answers/photos
		Minimum	Partial	Full	
Response format					
Conventional	892	100	100	98.9	11
Image-based	345	39.7			2.9
Group					
Text	632	100	100	98.6	11
TextChoice	260	100	100	99.6	11
Images	302	**37.7**			2.9
ImagesChoice	43	**53.5**			2.7

Note: bold notes statistically significant differences among formats, and among format-corresponding groups.

Online First publication date: 13-11-2024

Reflective Appendix

When collecting photos of books at home from participants in an online panel, who are predominantly accustomed to answering conventional survey questions, several challenges arose. First, there were difficulties in study design, including decisions about which information to request, what to exclude, whom to target, and how to organize the classification process. Some challenges were related to the substantive questions (e.g., asking for children's grades) and others to the collection of images (e.g., whether to ask for screenshots).

Second, limitations emerged during the data collection and processing stages: only a small number of respondents preferred photos over the conventional format, and more than half did not submit photos when requested due to factors such as privacy concerns, technical issues, or not being at home. Furthermore, some of the photos submitted lacked key information, and the manual classification process led to inconsistencies across researchers, delaying the data analysis.

Challenges in the design phase were anticipated and addressed in advance. However, those that emerged during the data collection and processing stages proved more complex. Notably, fewer participants than expected chose for and submitted photos, which limited the ability to conduct certain supplementary analyses. While these additional analyses were not central to the study, they would have provided valuable insights. As a result, the primary analyses presented in the main paper were completed as planned, but greater participation and preference for photo submissions would have allowed for a more comprehensive exploration of the data.

These issues are likely to persist as long as photos remain an emerging and unfamiliar data type for survey respondents. Consequently, researchers should anticipate facing similar challenges in future studies. However, these obstacles can potentially be mitigated by implementing the recommendations outlined in the final section of this appendix.

Design Difficulties

The first difficulty faced when designing this survey was the definition of the sample. Initially, since this project was designed in collaboration with substantive researchers who wanted to study the link between the books at home and the children's grades at school, the target population of interest were parents of children in primary school. However, changes in the regulation in Spain regarding the evaluation system (see section "Sample and data collection" in the paper) made it necessary to adjust the target population to parents of children in first, third, or fifth year of primary school. Without this adjustment to the sample, respondents might not be able to provide a grade depending on the year of pri-

mary school attended. This adjustment led to a lower quantity of Netquest panelists that matched the required survey profile. Further, a 3% error margin was added for the quotas to consider the lack of exact information about the target population and ensure their fulfillment.

Second, respondents in the conventional format were asked about the total number of books and their distribution among the following categories: books for toddlers and children who do not know how to read, books for literate children and teenagers, and books aimed at a general audience. While the need for including categories was clear, it was difficult to decide which exact categories to use and how to classify books within these categories (e.g., books containing text and drawings could be intended for toddlers, read to them by others, or for literate children capable of reading on their own).

Third, one initial concern regarded eBooks and the potential difficulties to capture this information. If asked for photos, respondents might have had to photograph the screen of the device, which might or not have contained the cover, titles, or author of the books. Similarly, if they were to answer the survey from the same device used for reading, respondents would have had to quit the survey, take screenshots, return to the survey, and then upload them. Ultimately, the collection of eBooks was discarded given results in previous literature indicating their lack of impact on children's academic achievement and cultural capital (Heppt et al., 2022; Otte, 2023; Pagel & Heppt, 2016).

Fourth, photos were first classified manually for two main reasons: 1) technical limitations, as discussions with computer vision experts revealed that algorithms needed improvement to accurately count books in photos when arranged in various ways, thus having a training dataset would be helpful, and 2) enhancing the algorithms was beyond the team's skills. Therefore, manual classification was the most fitting option to start for the project. However, the research team always considered that implementing an automatic classification, and comparing it with the manual one, would also be of high interest. Thus, collaborations were established to further investigate this option.

Main Limitations During Data Collection and Processing

One of the limitations was the inability to control whether respondents were answering from home, which was relevant for both response formats: being at home could help provide a better estimate for conventional questions, and it was essential for sending photos of the books at home. Initially, the inclusion of a question asking where respondents were answering the survey was considered (e.g., Revilla and Couper, 2021, did that when asking for voice answers), but the location could not be confirmed as geolocated data were not collected. Consequently, such a question would not distinguish between participants who were actually at home from those merely stating they were. Instead, a reminder was included at the survey's outset, urging participants to respond from their home

locations (if they were not at home, they could leave and re-access the survey with the link included in the invitation e-mail or a link available in the fieldwork company's app). Regrettably, this instruction seemed to go unnoticed or was disregarded by some respondents when accessing the survey. Indeed, 8% of respondents who did not submit photos when asked, stated that they were not at home. However, this number could be higher since "not being at home" was not an option in the radio-button question that asked for reasons for not uploading photos. Instead, respondents wrote this reason in the "Another reason" category, where they could type their own explanations. I expect that not being at home might have impacted the quantity of photos submitted, although it is uncertain whether those respondents would have submitted the photos even if they had been at home.

Additionally, fewer people than expected chose the images-based format (only 4%). Although H1 stated that respondents would prefer the conventional format over the images-based one due to the complexity of the task, I did not expect the difference to be so pronounced as previous research, based on a simpler task, found a higher preference for photos (Ilic et al., 2022). Moreover, around 40% of respondents actually asked for photos shared at least one. The low resulting proportion in the preference for photos and participation when asked for them, prevented the execution of some of the planned analyses, as assessing the impact of the respondents' characteristics, behaviors, and opinions on their preference for one response format over the other. Further, there are less images remaining for the analyses on data quality.

Moreover, photos were not always clear. Each dwelling is different, and while some books are well organized and clear in the photos, others are arranged in ways that is not possible to read the title or discern features that would allow classification in the three categories. This is critical especially for children's and teenagers' books, since these books are often stored in ways that make it more difficult to extract the information. This difficulty was anticipated when designing the survey, and thus specific instructions (which are available in the protocol by Iglesias et al., 2023) were designed to try to minimize the problem. These instructions aimed to include all the relevant information while also being as concise as possible. They explained how the photos should be taken in terms of lightning, distance to the books, and exclusion of distracting items. However, it was not expected that respondents will follow all the instructions since a lot of efforts on their side was needed to do so (e.g., removing several personal items from in front of the books). Moreover, the instructions and visual examples shown were related to adults' books in shelves: instructions for children's books or other types of storage were not presented, even if they might have been relevant.

Further, since classification was conducted manually, inconsistencies between researchers due to the complexity of the task were identified, especially in assigning books to categories. This was addressed by constantly review-

ing differences between classifiers and re-classifying photos (thus extending the duration of the classification step), but inconsistencies were not completely eliminated. Fewer challenges arose regarding consistency in storage and languages, but errors might still have occurred due to the coexistence of different languages in Spain and the fact that researchers were not fluent in all of them. In any case, the photos and manual classification outcomes are intended to be used by computer vision experts to improve existing algorithms. These improvements are expected to facilitate at least an accurate total count of books, which would be a significant contribution for research involving the collection of photos of books. The potential of such improvements will depend on both the resources available to the computer vision experts and the accuracy of the initial manual classification, since errors in manual classification of the training photos (i.e., photos of books collected and classified in this study) might lead to inaccurate results of the algorithms.

Recommendations for Future Research

Based on the experience in this study, the recommendations for researchers are:

1. To plan ahead when considering the collection of photos. Many extra steps are needed, starting with the programming of a tool allowing the collection of photos in a survey, which needs to be tested in different moments, operating systems, and browsers. These tests might lead to time-consuming improvements. Furthermore, images need to be stored in safe folders with enough capacity as to contain all the photos, which can be of several megabytes. Further, the servers storing such folders need to operate quickly to ensure that the experience of respondents uploading the photos is positive. Servers offering this service are for payment, thus funding for this needs to be assigned when planning the project.

2. Clear operationalization of the items to be observed in the photos is crucial from the outset, as this could be the first step enabling researchers to discern whether images are the best fit for their study. In this regard, researchers must clearly define the items they want to extract from the images, and establish the method for extracting the information. For more information on this matter and the other steps to be considered before, during, and after image collection, see Iglesias et al. (2024).

3. The definition of the items should also be conveyed to respondents when it does not interfere with the project's objective, so they can easily identify if the items of interest (e.g., books) are clearly visible in the photos.

4. Finally, the classification of the images is a critical issue. In any research, the method of classification (manual and/or automatic) needs to be defined, and the necessary resources must be allocated accordingly. In the case of this study, classification was manual. For this purpose, guidelines and examples were created to train the classifiers. However, it is important to note that manual classification is time-intensive, demanding meticulous attention to details and potential problems with the photos.

Researchers interested in collecting photos through surveys should be aware that there are numerous practical challenges involved in the design, collection, and analysis stages, more than with conventional questions alone. Therefore, it is crucial for researchers to have a well-defined plan for the entire process to ensure that photos are collected and analyzed successfully, making them valuable in addressing the research questions.

To achieve this, the guidelines provided by Iglesias et al. (2024) could be particularly helpful, as they offer a comprehensive overview of the entire process, from operationalization to analysis. However, given that challenges may still arise, researchers should remain flexible and be prepared to adapt their approach as necessary. For example, they should anticipate potential low participation rates and have a contingency plan in place, such as supplementing photo collection with conventional survey questions for respondents who do not provide photos.