Unobtrusive measures for media research

Araujo, T.; Neijens, P.

DOI
10.1002/9781119011071.iemp0049

Publication date
2021

Document Version
Final published version

Published in
The International Encyclopedia of Media Psychology

Citation for published version (APA):
Araujo, T., & Neijens, P. (2021). Unobtrusive measures for media research. In J. van den Bulck (Ed.), *The International Encyclopedia of Media Psychology* (Vol. 3). (The Wiley Blackwell-ICA International Encyclopedias of Communication). Wiley Blackwell. https://doi.org/10.1002/9781119011071.iemp0049

Unobtrusive Measures for Media Research

THEO ARAUJO
PETER NEIJENS
University of Amsterdam, The Netherlands

Often employed measures in media research are typically based on self-reports articulated in surveys, interviews, and focus groups. These reactive or *obtrusive measures* run the risk of being invalid as the research context may affect the responses obtained. Several aspects of the reactive measurement process may contribute to the invalidity of reactive measures. Webb, Campbell, Schwartz, and Sechrest (1966) distinguish respondent and researcher factors. The first category includes social approval and social desirability when answering questions, response sets, and role selection. Role selection is a validity issue as "validity decreases as the role assumed in the research setting varies from the usual role present in comparable behavior beyond the research setting" (Webb et al., 1966, p. 16). The second category includes interviewer effects, changes in the research instrument across research subjects and time periods, and sampling issues.

Webb et al. (1966) is usually considered the first publication that systematically introduces another type of measure: *unobtrusive measures*. This landmark book contains many creative, funny, and interesting examples and a thorough discussion of their strengths and weaknesses. Earlier interest in this type of measures was the basis for the Mass Observation movement (founded in 1937) which focused on observation and diaries to "create an 'anthropology of ourselves'" (Mass Observation, n.d.). Even earlier (1924), the well-known Hawthorne studies by Elton Mayo were among the first to demonstrate the reactive effects of research.

This entry gives an overview of the possibilities and limitations of different types of unobtrusive measures, with a focus on the online context in general, and social media in particular.

(Un)obtrusive measures: a typology

Nonreactive or unobtrusive measures are measures that do not intrude on or interfere with the context of the research (Trochim, n.d.). Or, according to Webb et al. (1966, p. 2), "measures that do not require the cooperation of a respondent and that do not themselves contaminate the response." "Unobtrusive measures are based on observing rather than gathering information directly from participants" (Connelly, 2017, p. 59). Typical examples of unobtrusive methods and data are observation, physical traces, and archives. Unobtrusive measures of behavior and language have been used as indicators

The International Encyclopedia of Media Psychology. Jan Van den Bulck (Editor-in-Chief), David R. Ewoldsen, Marie-Louise Mares, and Erica Scharrer (Associate Editors).

DOI: 10.1002/9781119011071.iemp0049

for behavior as well as underlying psychological constructs like values, personality traits, and so on (Hill, White, & Wallace, 2014; Webb et al., 1966). The emergence of online communication environments and the digitalization of physical environments have increased the possibilities for unobtrusive observation of attitudes and behavior.

The boundary between reactive and nonreactive research is not absolute. On the one hand, for example, in traditional "reactive" surveys and experiments researchers may use implicit measures such as eye tracking, registration of heart rate, and other physiological measures. In these cases, research subjects might be aware that they participate in a study, but they may be less aware of, or may have less control over, their (physiological) responses. Answers to questions in surveys or interviews can also be less obtrusive—for instance, less influenced by social desirability and social approval—if research participants are less aware of the intention that a researcher has for their questions, or if participants are allowed to answer privately, as in online questionnaires. On the other hand, unobtrusive measures may also be influenced by some kind of social desirability or social approval. In the case of social media, for example, users engage in strategic self-disclosure and self-presentation activities (e.g., Krämer & Winter, 2008) toward "imagined audiences" (Marwick & boyd, 2011). Based on the above, a typology of obtrusiveness of measures can be proposed as shown in Table 1. The upper left cell in the table shows examples of typical "obtrusive" research situations, and the bottom right examples of "unobtrusive measures." The two other cells are less typical.

Traditional unobtrusive measures often measure the behavior of a population at the aggregate level, including physical traces (e.g., sales data, accident rates) or observation of public behavior, and usually lack information about the individual characteristics of those observed. The online context offers more possibilities for unobtrusive measurement of individual behavior. These are discussed below.

Nondigital unobtrusive measures

The variety of unobtrusive measures is enormous as are the ways in which they can be applied. Three main categories are typically discussed.

Table 1 A typology of obtrusiveness.

	Subject participates in a study	*Subject does not participate in a study*
Subject is aware of a researcher or a public	Questionnaire Interview	Social media activities (e.g., tweets/posts, comments, likes, shares)
Subject is not/less aware of (the intention of) a researcher or a public	Answers to researcher's questions given privately Implicit measures (e.g., psycho-physiological indicators)	Physical traces (e.g., number of accidents, search terms) Observation (e.g., physical distance) Documents (e.g., diary)

Observation (of behavior or language) includes indicators based on the number and nature of social contacts, distance of subjects, seating arrangements, body positions, facial expressions, and use of specific words. Observation can take different forms depending on how the observer participates in the process. A distinction can be made between an intervening versus nonintervening observer (thus in naturally occurring settings); between participant versus nonparticipant observation; between visible (overt) versus nonvisible participation; between observation with versus without knowledge of the subject; and between observations with or without (hidden) hardware (e.g., microphone, camera).

A common distinction in *physical trace*s is between erosion "where the degree of selective wear on some material yields the measure" and accretion "where the research evidence is some deposit of materials" (Webb et al., 1966, p. 36). Examples of the first category include wear level of book or magazine pages, or of floors, and theft patterns in hotels. The second category comprises indicators such as clothing, litter, fingerprints, and graffiti. These measures can be acquired in an uncontrolled (or naturalistic) context or in a controlled (or contrived) context as part of an experiment. Physical data are "best suited for measures of incidence, frequency, attendance, and the like" (Webb et al., 1966, p. 46).

Archives and *documents* include "archival data from records and documents related to activities of people, institutions, and other groups" (Connelly, 2017, p. 59), such as sales, actuarial, institutional, or voting records, directories, annuals, yearbooks, minutes, mass media, and personal documents (such as diaries and letters). For example, advertising images are used as indicators of social change, and accident rates as an indicator for success of a traffic measure.

Digital unobtrusive measures

The emergence and increasing prevalence of online communication environments and the subsequent digitalization of offline environments have opened up a series of opportunities for measuring and tracking human behavior at an unprecedented scale. Behavioral tracking focuses on the digital traces that a user leaves behind. These include visits to websites, search behavior, likes and shares of posts, comments, reviews, and online buying. Offline user interactions, such as confrontations with out-of-home advertising, visits to physical locations, and individual activities (e.g., purchases in stores, routes taken in public spaces, etc.) are also increasingly captured through Wifi and Bluetooth grabbers, beacons, the registration of activities of "tagged" devices, and via cameras. Behavioral tracking makes it possible to trace the user's journey along the various (digital) touch points and gives "real-time" information about their attention, engagement, interest, and (buying) behavior.

This tracking can take place at different levels. On the one hand, patterns of search activities at aggregate levels have been proposed as a proxy to forecast economic or consumer activity (Choi & Varian, 2012) with varying levels of success. Likewise, patterns of website visits (e.g., pageviews, unique visitors, average time spent on page, most common paths) can be used to observe patterns of user behavior and options at aggregate levels. A common approach is to collect electronic word-of-mouth (eWOM) or, more broadly, user-generated content on platforms such as review websites, or online communities. Researchers can also connect to social media application programming interfaces (APIs) and download publicly available data with applications created for research. This form of data collection is, however, increasingly restricted by social media platforms, with scholars warning that researchers need to be ready for a "post-API age" (Freelon, 2018).

On the other hand, digital behavioral tracking can focus on individuals, be it by the creation of user profiles via cookies generally placed via advertisement and website visits or the actual tracking of specific individuals. For the latter option, researchers make use of tracking software installed in the devices of participants in a study (Araujo, Wonneberger, Neijens, & De Vreese, 2017). This type of software is generally able to collect information about the participant's online activity (e.g., websites visited, mobile applications used), with several limitations and restrictions that need to be considered, especially for activity on smartphones (for an overview or examples, see Araujo et al., 2017; Boase & Ling, 2013). Similar approaches include partnering with participants to have them donate their own data, including their navigation data via specific browser plugins (e.g., Menchen-Trevino, 2016), their social media activity via applications created to collect data through social media APIs, or by having users export and donate their own social media data using their rights to data portability (e.g., Halavais, 2019).

Benefits and limitations of unobtrusive measures

An obvious advantage of unobtrusive measures is that they give direct access to behavior and opinions not influenced by the research situation. Furthermore, unobtrusive measures make it possible to measure behavior of groups in society which are hard to reach, ranging from busy CEOs of companies to nonmainstream communities; may outperform traditional measures in studies on sensitive topics; and allow collecting data from older or for longer time periods. Digital data, in principle, come in great numbers, measure a user's actions in real time, can be cheap and require little effort, and give a detailed picture of a user's online behavior. For example, behavioral tracking gives campaigners a fast and simple opportunity to test various variants of advertising, political messages, and other promotional online strategies, a practice known as A/B testing or multivariate testing.

Unobtrusive measures have limitations as well. The most obvious issue is *validity*. Archives run the risks of "selective deposit" (how selective was the inclusion of the documents?) and "selective survival" (how selective was the preservation of the documents?) (Webb et al., 1966, p. 54), and changes in definitions across time. Physical traces may have been shaped by other than the theoretical construct. Observation measures may suffer from the presence of the observer (more probable as the observer intervenes more strongly in the situation), and the ability of humans to observe, memorize, and register "what is seen and heard." Digital data may be manipulated by bots "visiting" websites or social media platforms in order to artificially raise the number of visitors or engagement with posts. As digital data are often data about devices rather than about individuals, issues may arise for the identification and tracking of individuals, not only because they use multiple devices (smartphones, tablets, laptops, home and work computers), but also because some of these devices are also used by others.

Validity is especially questionable when behavior-based unobtrusive measures are used for assessing underlying psychological concepts. Hill, White, and Wallace (2014) illustrate this by referring to different studies in which words such as "my," "mine," "I," and "me" are used to measure completely different concepts such as Machiavellianism, narcissism, and self-confidence (p. 158). And what does "liking" of a message mean? Does a user really like the message, or do they want to please the messenger (boyd & Crawford, 2012)? Hill, Kern, and White (2014) studied the lack of validity of seven existing unobtrusive measures of "executive overconfidence." They recommend that research should provide "a rationale connecting the measure to the construct that the measure is attempting to assess while simultaneously ruling out other constructs" (p. 158).

Privacy issues play an important role. Individuals often do not know that "their" data are used to make decisions about them, nor how, and questions about data persistency, repurposing, and spillovers (i.e., data produced by third parties about an individual) arise in the privacy debate (Tucker, 2019). Concerns about privacy have led to new policy measures restricting the collection, storage, and use of data: in the European Union, the General Data Protection Regulation has come into effect and in the United States policy recommendations and changes continue to follow suit.

Another issue is lack of *representativity*. Not everyone is online, not everyone online expresses themselves, and not everyone online who expresses themselves does so in public content that can be accessed by researchers. A case in point is the usage of Twitter as a proxy for public opinion (as discussed in, e.g., boyd & Crawford, 2012): Twitter users do not represent the overall population, the tweets posted do not represent the behavior of all Twitter users (e.g., those who do not post), and even user activity on Twitter is not representative of activities that users may have in more private social media platforms (e.g., Facebook) or encrypted messaging environments (e.g., WhatsApp) (see the role selection issue mentioned above). Caution about the representativity and generalizability of the claims is strongly advised.

Unobtrusive data are often *unstructured*. They may include website visits, liking of messages, comments on YouTube videos, photos, and videos. Automatic coding of texts including machine learning-based procedures is used to code the content of these communications, but that may come with serious measurement errors. Dross rate ("the part of the conversation that is irrelevant to the topic at hand"; Webb et al., 1966, p. 32) may affect the efficiency of unobtrusive measures. Finally, as indicated above, unobtrusive measures often lack profile information about the user, such as demographics and other characteristics.

All in all, unobtrusive measures add to the toolkit of the researcher, and, as with all measures, they have strengths and weaknesses. They can be used as an add-on to reactive methods such as questionnaires, interviews, and focus groups to get a fuller picture of the phenomenon being studied, or to provide triangulation (Connelly, 2017).

As many others have voiced, "all methods are fallible and methods that are fallible in different ways complement each other even if one is absolutely weaker than the other. Convergence of findings by two methods with different weaknesses enhances our belief that the results are valid and not a methodological artifact" (Bouchard, 1976, p. 267).

SEE ALSO: Attitude; Big Data, Analysis of; Big Data, Collection of (Social Media, Harvesting); Ethics of Empirical Research; Experiment, Other: Natural, Field; Measurement of Attitudes; Measurement of Media Exposure; Measuring Behavior in Media Psychology; Observational Methods; Survey Methods, Traditional, and Public Opinion Polling; Validity; Wearable Devices for Health-Related Data

References

Araujo, T., Wonneberger, A., Neijens, P., & De Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of Internet use. *Communication Methods and Measures*, *11*(3), 173–190. doi:10.1080/19312458.2017.1317337

Boase, J., & Ling, R. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*, *18*(4), 508–519. doi:10.1111/jcc4.12021

Bouchard, T. J. (1976). Unobtrusive measures. *Sociological Methods & Research*, *4*(3), 267–300. doi:10.1177/004912417600400301

boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679. doi:10.1080/1369118X.2012.678878

Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, *88*(s1), 2–9. doi:10.1111/j.1475-4932.2012.00809.x

Connelly, L. M. (2017). Unobtrusive methods. *Medsurg Nursing*, *26*(1), 59–61.

Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, *35*(4), 665–668. doi:10.1080/10584609.2018.1477506

Halavais, A. (2019). Overcoming terms of service: A proposal for ethical distributed research. *Information, Communication & Society*, *22*(11), 1567–1581. doi:10.1080/1369118X.2019.1627386

Hill, A. D., Kern, D. A., & White, M. A. (2014). Are we overconfident in executive overconfidence research? An examination of the convergent and content validity of extant unobtrusive measures. *Journal of Business Research*, *67*, 1414–1420. doi:10.1016/j.jbusres.2013.08.011

Hill, A. D., White, M. A., & Wallace, J. C. (2014). Unobtrusive measurement of psychological constructs in organizational research. *Organizational Psychological Review*, *4*(2), 148–174. doi:10.1177/2041386613505613

Krämer, N. C., & Winter, S. (2008). Impression management 2.0: The relationship of self-esteem, extraversion, self-efficacy, and self-presentation within social networking sites. *Journal of Media Psychology*, *20*(3), 106–116. doi:10.1027/1864-1105.20.3.106

Marwick, A., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, *13*(1), 114–133. doi:10.1177/1461444810365313

Mass Observation. (n.d.). *History of Mass Observation*. Retrieved from http://www.massobs.org.uk/about/history-of-mo

Menchen-Trevino, E. (2016). *Web historian: Enabling multi-method and independent research with real-world web browsing history data*. Retrieved from https://www.ideals.illinois.edu/handle/2142/89388

Trochim, W. M. K. (n.d.). *Research methods knowledge base*. Retrieved from https://socialresearchmethods.net/kb/unobtrus.php

Tucker, C. (2019). Privacy, algorithms, and artificial intelligence. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence: An agenda* (pp. 423–437). Chicago, IL: University of Chicago Press.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago, IL: Rand McNally.

Further reading

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., … Van Alstyne, M. (2009). Life in the network: The coming age of computational social science. *Science*, *323*(5915), 721–723. doi:10.1126/science.1167742

Lee, R. M. (2015). Unobtrusive measures. *Oxford bibliographies*. Retrieved from https://www.oxfordbibliographies.com/view/document/obo-9780199846740/obo-9780199846740-0048.xml

Theo Araujo (PhD, University of Amsterdam) is an associate professor in the Amsterdam School of Communication Research (ASCoR) at the Communication Science Department of the University of Amsterdam, and codirector of the Digital Communication Methods Lab (digicomlab.eu) from the Research Priority Area Communication. His research focuses on the increasing adoption of artificial intelligence and related technologies within the communication environment, including conversational agents and automated decision-making, as well as on computational social science, with a special interest on large-scale digital data collection and analysis for communication research.

Peter Neijens (PhD, VU University Amsterdam) is Honorary Fellow and Emeritus Professor and Chair of Media and Persuasion in the Amsterdam School of Communication Research (ASCoR) at the Communication Science Department of the University of Amsterdam. His research interests include persuasive communication, in particular media uses and effects in advertising and public opinion. Peter is a Fellow of the International Communication Association (ICA), and received the Netherlands–Flanders Communication Association (NeFCA) Career Award for a lifetime of scholarly achievement in communication science, as well as the European Advertising Academy's (EAA) Flemming Hansen Award for long-term impact in advertising research.