

Kay Brauer and Rebekka Sendatzki

The Interview as an Assessment Method in Psychology

Abstract: The interview has a tradition in psychology. In this chapter, we discuss it as an assessment method within the field of psychology, first discussing the general measurement problem in psychology (i.e., assessing non-observable constructs). We then give an overview of findings on the objectivity, reliability, and validity of interviews, characteristics that distinguish different types of interviews, their merits and disadvantages (e.g., in comparison to other assessment methods such as subjective self-reports in questionnaires), and their usage in applied settings such as personnel selection and clinical assessment. In conclusion, we posit that the interview remains an important method to generate data to derive diagnostic information in psychology.

Keywords: assessment, diagnostic, job interview, objectivity, psychology, qualitative, reliability, standardization, validity

When laypeople think of psychologists, they typically think of the stereotype of a psychotherapist or psychoanalyst in the tradition of Sigmund Freud, who interviews and speaks with patients (e.g., BDP; Jiménez and Raab). Thus, psychology is partly synonymously linked with the interview as a method to collect information and to treat clients and patients. However, the understanding of psychology has transformed into an empirical science in the tradition of natural sciences. Thus, there is a strong emphasis placed on the objectivity, reliability, and validity of the methods used to generate data. While modern psychological science uses interviews to generate data less frequently in comparison to its beginnings, interview techniques remain a powerful tool of psychological assessment in research and applied disciplines, for example, in both the clinical assessment of psychological disorders and personnel selection. In this chapter, we will give an overview of different types of interviews and discuss their usage, merits, and limitations from the viewpoint of psychological assessment. Before discussing the interview as a concept in more depth, we first give a short overview regarding the change in data generation within psychology over time and the challenge of assessing psychological characteristics. This might help readers unfamiliar with assessment methods in psychology to understand the *general* aims and problems of psychological assessments. We will discuss interviews from the perspective of their objectivity, reliability, and validity, and illustrate their usage in examples

of personnel selection and clinical assessment. Finally, we compare the interview method with the frequently used questionnaire method.

Psychology as a Science

As mentioned, the role of the interview in psychology has changed in the last century, whereas quantitative data collection and analysis methods have received increased interest. To understand this trend, one has to investigate the history of psychology as a science throughout the last century. Modern psychology is an *empirical* science that aims to describe, explain, and predict cognitive and affective processes and behaviors (see Zimbardo et al.). Following the distinction between quantitative (i.e., using quantifiable and scalable units) and qualitative (i.e., knowledge derived from “soft” data sources such as interviews) approaches, psychology has been a quantitatively driven science since the late nineteenth century, beginning with the works of Wilhelm Wundt, William James, Hermann Ebbinghaus, and others on human perceptual processes, short- and long-term memory, and learning processes (for an overview see, for example, Mandler; Mischel). They began to collect and analyze quantitative data on individuals’ reactions (e.g., reaction times, frequencies of remembered stimuli). In the early twentieth century, quantitative data began to dominate psychological research (see, for example, Young). Similarly, the development and increased use of self-report instruments (i.e., questionnaires) to assess individual differences in personality traits, values, and attitudes, as well as the availability of complex data analysis methods such as correlation and factor analysis in the early twentieth century, led to psychology transforming into a data-driven natural science (see, for example, Bollmann; Vincent; Young). Thus, only minor space was left in psychological research for “soft” assessment techniques such as interviews.

At the same time, psychoanalytic works stood in contrast to the aforementioned quantitative strategies but, nevertheless, received great attention both within and outside of psychology. In contrast to quantitative approaches to psychology, psychoanalytic theories were frequently theory-driven and rarely systematically tested empirically (e.g., Freud’s theory of personality development; for a discussion see, for example, Fisher and Greenberg). This had implications for the role of the interview in psychology: the underlying notion of psychoanalysis is that psychological characteristics and processes are not directly accessible but must be “uncovered” through *talk therapy* in which the psychoanalyst gains insight into the client’s feelings and thoughts. This view generally echoes the approach of psychological assessment aiming to measure latent traits (see the

following section). While the methodological approaches to assessing psychological processes, states, and traits have been extended in recent decades (e.g., Ecological Momentary Assessment, Daily Diary Methods, and Smartphone Sensing; see, for example, Harari et al.; Smyth and Stone; Wu and Clark), the interview technique remains an important method of psychological assessment. Although it might seem that quantitative and qualitative approaches to psychological assessment are diametrical opposites, it must be noted that they share the same aim, namely, assessing non-observable psychological constructs that describe the experiential world of individuals.

The Challenge in Assessing Psychological Constructs

To understand the role of the interview for psychological assessment, it is crucial to clarify the general challenge of “measuring” psychological constructs and how they are expressed in affect, thought, and behavior. In contrast to physics or chemistry, where variables such as the temperature, weight, or height of objects can be observed and *measured* in the narrow sense (e.g., by using an objective, reliable, and valid measurement instrument such as thermometers, scales, or rulers), psychological characteristics such as personality traits or intellectual abilities (e.g., intelligence or attention) cannot be directly *measured* as they are not directly observable (see, for example, Michell). For illustrative purposes, one might imagine assessing the expression of a personality trait: for example, extraversion is characterized by enjoying human interaction and is expressed by behaviors such as being talkative, assertive, and sociable (e.g., Costa and McCrae). It is not possible to measure the expression of one’s extraversion in a similar way as in physics because the characteristic of interest is a *latent* construct—no scaling device or ruler for extraversion exists. To address this measurement problem, psychologists have to approximate the latent non-observable trait by collecting information about manifest indicators (i.e., those observable to oneself or others) that allow inferences about the expression of the underlying latent trait. Of course, such indicators should be valid and correlate with the trait of interest discriminately (i.e., with no other traits) to allow robust conclusions. When considering our example of assessing extraversion, one might be interested in responses to indicators such as “do they speak loudly” or “do they often attend parties,” or “do they like talking to people” or subjective descriptions with adjectives such as “affectionate” in contrast to “reserved,” “talkative” in contrast to “quiet,” or

“joiner” in contrast to “loner” to infer one’s level of extraversion (e.g., McCrae and Costa). We would conclude that someone is highly extraverted when they endorse such statements that indicate extraversion or describe themselves with the named adjectives.

Taken together, the main aim of psychological assessment is to provide and analyze indicators that reliably, objectively, and validly operationalize a given latent construct (e.g., personality traits, cognitive abilities, or creativity, to name but a few). The evaluation of such indicators that allow conclusions about latent constructs to be drawn might be realized either by quantitative (e.g., using self-report questionnaires in which responses are quantified to scores that reflect expressions of psychological traits) and/or qualitative approaches, with the interview belonging to the latter category. While both approaches should ideally lead to the same conclusions, they differ regarding a number of formal factors that have consequences for the data analysis and are discussed in later sections of this chapter. After introducing the three main criteria that must be met by any assessment technique to derive robust information, we will discuss the merits and issues with the interview method from the point of view of these criteria.

Objectivity means that results are independent of the researcher (or interviewer) and that no other contextual variables (e.g., confounders such as noise, the interviewer’s mood, or the weather) affect the assessment process. For example, a personnel selection interview is considered to be objective when fixed criteria based on a priori selected points concerning the aim of the interview (e.g., what information has to be collected concerning each candidate) are used to construct it. In our example of personnel selection, independent interviewers should use the same questions to derive information about the applicants’ suitability for the open post, and the interview questions or addressed topics should not depend on the interviewer. Moreover, information should be assessed in the same way across participants (i.e., by using the same type of questions with minimal deviations in the wording thereof). Applying high standardization by using a-priori criteria and clear strategies for assessing them should result in a selection process that is only minimally dependent on the interviewer or other external sources (for a discussion see, for example, Latham and Saari). However, interviews differ with regard to their structuredness, as we will discuss later in the “Types of Interviews” section.

Reliability is characterized by the consistency and accuracy of the assessment method. The results and information obtained from a reliable experiment, questionnaire, or interview should be identical, or at least very similar, in repeated *ceteris paribus* measurements. For example, an interview would be

considered reliable when the derived information leads to the same conclusions when conducted repeatedly with the same respondent.¹

Validity means that the assessment instrument (e.g., questionnaires and interviews) assesses what it claims and aims to assess; for example, by collecting information that is distinctively indicative of the latent trait of interest. For example, the question “how often have you had a drink containing alcohol during the past week” is a valid indicator for examining drinking habits as it allows conclusions concerning the underlying latent construct of “substance use.” Assessing and evaluating these three criteria for interview methods is difficult because the qualitative nature of the collected data does not permit the use of the standard quantitative approaches to evaluate reliability and validity (e.g., computing internal consistency as an indicator of reliability; using factor analysis for validity analyses etc.; e.g., Furr and Bacharach). However, the past few decades have seen efforts to evaluate the objectivity, reliability, and validity of interviews by aggregating findings across studies (meta-analyses) and providing a database for such analyses.

Objectivity, Reliability, and Validity of Interviews

First, it must be noted that a singular interview method does not exist as the term describes a classification of techniques that differ in their structure, approach, and course depending on the context and aim of the interview technique. When we use the term “interview” without further specification, we mean the minimal definition of what it constitutes; namely, the interaction between one person who aims to collect information (also called “assessment”; i.e., the interviewer) and someone who provides information to the interviewer (i.e., the interviewee). As we will discuss later, interviews differ regarding numerous characteristics such as the level of standardization (i.e., non, semi, or fully structured; see “Types of Interviews”) and context (e.g., clinically oriented; job interview; interview for research purposes) that also play a role for its objectivity, reliability, and validity. In line with this minimal definition, we will discuss the factors that play a role during the interview process in this section.

¹ The notion of repeated measurements is a theoretical illustration. Of course, one would expect that changes over time can occur; for example, when comparing the interview of a patient with depressive symptoms before and after successful psychotherapeutic and/or psychopharmacological interventions. In this case, one would expect that changes in depressiveness would be detected by the interviewer.

Objectivity. The interview approach is affected by subjectivity in both the interviewer as the receiver and processor of information and the interviewee (or respondent) as the source and provider of information (e.g., Morrison; Nordgaard et al.). As displayed in figure 1, an interviewee's responses are affected by their perceptions and interpretations of the true facts. When retrieving such information internally, the interviewee reports their recounted information to the interviewer. Taking the subjective nature of this process into account, how reliable and valid the interviewee's reported information is in relation to the factual event cannot be measured objectively.

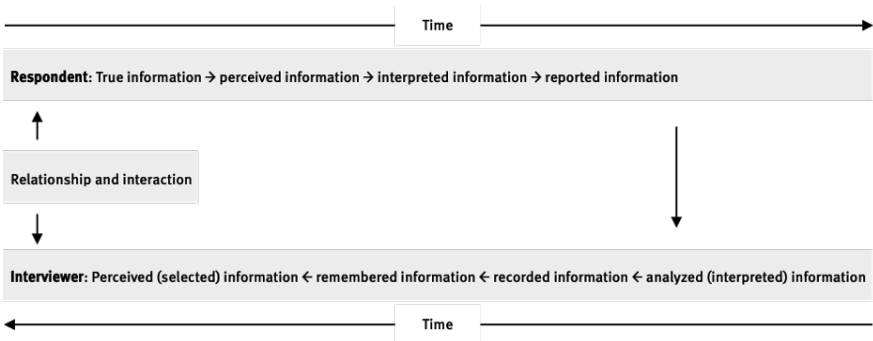


Fig. 1: Process and structure of interviews concerning the respondent, interviewer, and the unique dyadic interaction between interviewer and respondent.

Further, the interviewer needs to decode the reported information provided by the interviewee, which is also affected by subjective interpretations concerning how the interviewer perceives the reported events, which content the interviewer remembers, and which details are recorded. Such records are typically based on memory logs or audio and/or video records. Thus, the interviewer, who assumes a subjective role in the interview process, is also a source of reliability and validity concerning the diagnostic information collected. Considering the amount of subjectivity in both interview partners, one can conclude that objectivity is not perfectly displayed in the interview method. However, objectivity can be increased through standardization; for example, by using fully structured interviews instead of non-structured interviews and training interviewers in how to conduct interviews according to certain protocols (e.g., Morrison; Nordgaard et al.). The use of structured interviews helps to guide the formal interview process and to follow rules to collect information in a standardized way. In addition, it minimizes individual decisions by the interviewer. An example of low objectivity would

be that an interviewer could be preconceived by potential foregone conclusions and may ask questions that fit into their hypothesis but are too narrow to assess the full background needed for a correct diagnosis² (e.g., Morrison; Nordgaard et al.). We will illustrate this point with an example from the field of forensic psychology: research on the accuracy of eyewitness testimonies and the reconstruction of memories has shown that the *wording* of a question already affects the retrieval of memorized events in respondents. In a classical experiment, Loftus and Palmer showed participants video footage of a car crash and subsequently asked them for their estimates of the speed at which the cars collided. In their question (“How fast were those cars when they...”), they systematically varied the verb (i.e., “contacted?”; “hit?”; “bumped?”; “collided?”; and “smashed?”) and tested whether this might affect people’s estimates of the crashing cars’ speed. As a result, they found participants’ speed estimates varied as a function of the verb used, with “contacted” being related to the lowest speed and “smashed” being related to the highest speed estimate. The findings have been replicated well and show that how one is asked for information influences the response. This can be related to the interview process, as the interviewer’s questions can affect the responses by the interviewee—showing the influences of both the interviewee and the interviewer. Hence, it is important to standardize the content of interest and wording of questions to prevent selective questioning on the basis of interviewers’ preconceptions and to maximize comparability of responses among interviewees.

Beyond such subjective processes in both interview partners, it must be noted that the relationship and interaction between the interviewer and interviewee play a role in the reports given. There is robust evidence in the literature that dyadic interactions (i.e., the unique interaction between two persons) are denoted by the interdependence and characteristics of both dyad members who shape the interaction (for an overview and discussion, see Brauer and Proyer). For example, establishing rapport is important for creating an atmosphere that allows the respondent to talk about sensitive topics (e.g., Morrison), and studies have determined that establishing rapport increases respondents’ sense of interpersonal security and relates to respondents’ greater disclosure regarding sensitive topics (e.g., Henson et al.; Sun et al.). However, *how* rapport is established differs not only between interviewers but also depends on the respondent; it is thus unique to each dyad. This poses the issue that the interviewer has to balance two aims simultaneously: namely, acting in a way that ensures professional

² We use the term “diagnosis” in its broad definition of assessing a psychological phenomenon and not limited to diagnoses in the sense of identifying a clinical disorder.

distance and enhances objectivity, while at the same time establishing rapport by acting and reacting toward the respondent in a unique personal way to create an atmosphere that allows the respondent to speak openly—particularly when sensitive topics are the subject of the interview. There is no “formula” for establishing rapport, and the same rapport strategy can have different effects depending on the interviewee. For example, it is often suggested that smiling, as an expression of positive emotions, is fundamental to establishing rapport (e.g., Stocco et al.), but recent research has revealed that people differ in how they perceive being laughed and smiled at: a group of people experience smiling as malicious ridicule and a means of putting them down—independently of the intention and morphological characteristics of the smile (Ruch and Proyer). In the case of those people misinterpreting smiling (so-called *gelotophobes*; Greek: *gelos* = laughter, *phobos* = fear; Ruch and Proyer), using smiling to establish rapport can lead to participants quitting interview sessions, as gelotophobic respondents feel ridiculed by the interviewer (Platt et al.). Hence, while non-verbal behaviors such as smiling may increase rapport in the majority of interviewees, they can also have adverse effects. This example shows that the interviewer has to adjust the strategies of establishing rapport to the respondent’s reactions while simultaneously conducting the interview and aiming to collect diagnostic information. Hence, the dyadic relationship plays a role in the interview, and there is no one-strategy-fits-all formula to establish a good relationship between interviewer and interviewee. However, this unique dyadic interaction affects the objectivity and highlights the high cognitive demand interviewers face.

Finally, it must be noted that time, or more specifically, the time delay between forming and retrieving information (e.g., an episode depicting a prior event) from memory, affects the veridicality and biases of recollections (e.g., Lalande and Bonanno; Read and Connolly). Again, this affects the interviewer and interviewee alike, but it can be assumed that interviewees’ recollections suffer from greater time-related biases (e.g., when remembering episodes from childhood), whereas retrieval in interviewers is often comparatively short (e.g., covering the time span between interviews). However, using records (e.g., notes, audio/video records) often allows interviewers to address this issue and increase objectivity (e.g., Burnett et al.).

Taken together, objectivity is affected by psychological and formal factors that also affect the reliability and validity of interviews. To address the latter two, meta-analyses have helped to draw conclusions on reliability and validity.³ In a

³ Validity coefficients are described as correlations. Correlations range between -1.00 (perfect negative association between two variables) and 1.00 (perfect positive association), with 0.00

meta-analysis, the statistical coefficients of many independent studies addressing the same question are aggregated and statistically processed into an average coefficient. This approach allows the aggregation of knowledge across studies and has statistical advantages (e.g., higher statistical power than single studies). Two comprehensive meta-analyses on personnel selection interviews have advanced the understanding of the reliability and validity of interviews.

Reliability. Conway et al. aggregated 160 reliability coefficients from personnel selection interviews and found an average reliability of .70. They used the criterion of inter-rater reliability, which assesses the convergence between conclusions among independent interviewers (e.g., the consensus between interviewers in their decision to select a candidate), with higher convergence indicating higher reliability.⁴ This indicates that, on average, interviewers derive the same inferences based on their interview data. The coefficient meets the threshold for satisfying reliability (e.g., Furr and Bacharach). Further, Conway et al. investigated which factors contribute to the reliability of interviews and found that greater (a) standardization of interview questions (i.e., using pre-defined questions), (b) standardized interpretations of the responses, and (c) interviewer training were all associated with higher reliability. Using more advanced statistical techniques that allowed them to disentangle different sources of measurement error, Huffcutt et al.'s meta-analysis on job selection interviews again supported the notion that reliability increases with standardization and structuredness. Based on such findings, the literature recommends using structured interviews composed of a priori defined questions and on the basis of considerations concerning what should be assessed (e.g., psychological disorders; job-person fit; vocational interests etc.). Furthermore, training interviewers on how to conduct interviews (e.g., how to present themselves and react to interviewees) and interpret responses (e.g., using pre-defined criteria) improves their reliability (e.g., Craig; Latham and Saari; Rogers, *Diagnostic*, "Standardizing"; Wittchen). The objective of this training can also be achieved "naturally," as interviewers gain experience throughout their career. This contributes to understanding why trained clinical psychologists' diagnostic conclusions are comparatively reliable when conducting non-structured interviews; put simply, their years of experience in diagnosing patients support their judgmental processes and diagnostic inferences (e.g., Powell et al.). Additionally, psychometric theory

indicating independence between two variables. Reliability coefficients are interpreted similarly, with 0.00 indicating no reliability and 1.00 indicating perfect reliability.

⁴ More information on the statistical and theoretical background of this reliability approach can be found in Tinsley and Weiss' seminal paper.

shows that the reliability of judgments (e.g., on personnel selection or diagnoses) increases when using the judgments of more than one interviewer, as the biases and errors of individual interviewers are minimized by aggregating the interpretations and conclusions of multiple interviewers (e.g., Walker). Hence, it is recommended that interviews should be supplemented by an additional observer if the situation, context, and resources permit this decision.⁵ Findings on the reliability of *clinical* interviews are comparable, with a satisfying overlap among clinicians' diagnoses and the positive effects of training and structuredness (see, for example, Miller et al; Rogers, "Standardizing"; Widiger; Wittchen). Thus, available evidence suggests that interviews provide reliable information, especially when standardization is high.

Validity. Schmidt and Hunter's meta-analysis addressed the utility and validity of the interview approach in comparison to 18 other selection criteria (e.g., job experience in years, assessment centers, or reference checks, to name but a few) by testing the predictive values for the "job performance" and "training performance" outcomes in the field of personnel selection. They analyzed studies from 85 years of research and examined the associations between the performance in a selection interview and the measured job performance later in the job. Considering the time delay between the job interview and the assessment of performance, the correlations indicate how well the interview predicts subsequent performance. Schmidt and Hunter's findings identified that *structured* employment interviews predict job performance with a validity of $r = .51$, whereas *unstructured* interviews only reach a coefficient of $r = .38$. Overall, this ranks them second and ninth out of the 19 tested criteria, respectively. Further, the comparison of the coefficients (structured vs. unstructured) shows that the standardization of the interview plays an important role in the validity of the interview, with higher standardization going along with greater validity. They extended their findings by testing whether interviews contribute *beyond* the knowledge of applicants' intelligence, as measured by standardized cognitive mental abilities tests. When using intelligence as a baseline for predicting job performance, structured and unstructured interviews increase the validity of predictions, as they account for 24% (structured interviews) and 8% (unstructured) increases in validity. Thus, the findings again highlight the need for the standardization of interviews. Overall, Schmidt and Hunter's findings recognize that the interview is a useful and valid method in personnel psychology. Findings from the field of clinical psychology are also widely aligned with the evidence for validity, although it has

5 While it is typical that employment interviews are conducted with several interviewers/observers present, interviews in the clinical context are typically conducted by a single interviewer.

been criticized for its dearth of comparably large studies, such as that provided by Hunter and Schmidt, especially on the criterion of convergent validity (i.e., the agreement between interview and external data; see, for example, Miller et al.; Renner and Jacob; Widiger; Wittchen).

Overall, the findings from the literature show that the interview provides the means to collect information that facilitates the derivation of reliable and valid conclusions, increasingly so when standardization (e.g., structuredness, interviewer training) is high. As the findings also show, objectivity and standardization are important prerequisites for a reliable and valid assessment with the interview method.

Types of Interviews

As discussed, interviews differ regarding several criteria, such as their degree of standardization, which include structured, semi-structured, and unstructured interviews. This distinction regards their characteristics such as the questions asked, interviewee answers, data analysis (i.e., interpretation of the responses and methods to derive conclusions), and the interviewer's behavior (e.g., Craig; Miller et al.; Renner and Jacob). Which type of interview is used often depends on the field and diagnostic aim. Moreover, the training of the interviewer plays a role; for example, a well-trained interviewer who has conducted numerous diagnostic interviews over many years typically no longer needs structured or semi-structured interview templates as they have memorized the topics and questions that need to be addressed to obtain robust conclusions (e.g., Morrison).

Structured Interviews. In structured interviews, the questions are standardized, meaning that they are fixed in their number, wording, and order for each interviewee. This interview form enhances objectivity because each patient or client is interviewed under very similar circumstances, which allows comparability and ensures that variations in the given answers are caused by differences in the assessed trait and not by confounding variables or interviewer bias. For example, when a structured interview to assess depressive symptoms comprises the question "Have you been especially critical of yourself this past week, feeling you've done things wrong, or let others down?" (Williams), the question needs to be read out loud word-by-word by the interviewer. Further, it is clearly defined how to continue if the respondent affirms the question (i.e., if the interviewee responds with yes: "What have your thoughts been?") and how to record and interpret the responses ("0 = *absent*; 1 = *self-reproach, feels s/he has let people down*; 2 = *ideas of guilt or rumination over past errors or sinful deeds*; 3 = *present illness is a*

punishment. Delusions of guilt; 4 = hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations"; Williams). A crucial advantage of structured interviews is the minimization of bias and errors, but the consequence is that they are not as flexible and adaptive to the individual interviewee. In the course of the conversation, other aspects and topics may become more important than the ones the interviewer had prepared for. In a structured interview, open questions cannot be addressed in depth, which may cause a loss of information. Moreover, the structured and rigid form may seem too artificial to interviewees and in some cases not appropriate. For example, in a first clinical interview, patients should be allowed to talk freely about their circumstances, problems, feelings, and thoughts. This form is recommended as interviewees, particularly in the clinical sector, open up more easily if the interview feels more like a normal conversation where the interviewer (e.g., therapist) reacts to the addressed topics and asks further questions (e.g., Morrison).

Unstructured Interviews. In the unstructured interview, the purpose of the interview is fixed, and the topics and questions arise out of the situation, context, and conversation. Usually, this approach leads to a rich, detailed, and more individual conversation because the interviewee can respond more freely and add depth to their answers. Therefore, unstructured interviews can be described as more adaptive because the interviewer can advance into certain topics that are mentioned by the respondent and may be important for the course and outcome of the interview. Furthermore, follow-up questions can be asked depending on the interviewees' responses. The most fundamental limits of unstructured interviews are the missing comparability as well as less objectivity and reliability than in structured interviews (e.g., Conway et al.; Schmidt and Hunter). However, there is no guarantee that every important topic is discussed, as the conversation may go in a different direction than intended, which makes the interviewer's expertise even more important in ensuring a thorough assessment.

Semi-Structured Interviews. After introducing structured and unstructured interviews, one might think of a continuum where structured interviews are on one end of a pole and unstructured interviews on the other. The majority of interviews will not fall on either end, but will range somewhere on the continuum, making it semi-structured. This approach combines the advantages of both structured and unstructured interviews, allows for a certain degree of objectivity and flexibility, and widely avoids their disadvantages. Semi-structured interviews are based on a fixed structure to a certain extent but can be individualized and adapted situationally to acknowledge interindividual differences in the interviewees (e.g., Renner and Jacobi). This is realized by using a pre-defined catalogue of broader topics and questions (e.g., "examine depressive mood" or

“check for depressive symptoms”) instead of using a catalogue of previously formulated questions word-for-word, as shown in the example of structured interviews. The interviewer might make a list of topics and some questions beforehand to remember the most important aspects but use these as an orientation rather than a fixed guideline. Note that semi-structured interviews are particularly effective when carried out by interviewers with a certain degree of training and experience that allows them to adapt the interview individually.

The decision to use structured, unstructured, or semi-structured interviews depends on the aim of the interview. In certain situations, such as selection and diagnostic interviews, structured interviews should be preferred because objectivity, reliability, and comparability play a crucial role in these contexts, and errors based on interviewers should be kept to a minimum (e.g., Armoneit et al.). On the other hand, unstructured interviews are best suited for explorative settings, when the aim is to cover many diverse issues (e.g., Renner and Jacob).

Note that besides structuredness in conducting the interview, the *data analysis* can also range from structured to unstructured. There might be strict and predefined rules for analyzing and categorizing the given answers to compute a “score” for a latent trait (structured; e.g., Widiger; Williams), whereas the assessment of the given trait might also be decided solely according to the experience of the interviewer (unstructured). As with the standardization of conducting interviews, reliability increases when using standardized algorithms to interpret interviewees’ responses (e.g., by counting the occurrence of certain topics or symptoms; Morrison). Moreover, recent research has provided numerous digital and analogous approaches to analyzing qualitative data such as interviewees’ responses. For example, narrative analyses (McAdams; McAdams et al.) allow the identification of psychological themes by finding systematic patterns of content and topics in transcriptions of autobiographical interviews. Another approach is the quantitative language analysis of transcribed interviews. For example, the *Linguistic Inquiry and Word Count* (LIWC; for an overview, see Tausczik and Pennebaker) software scans interview transcriptions digitally and identifies keywords that are indicative of, for example, negative emotions (e.g., words that indicate anxiety, anger, or sadness). A merit of quantitative and structured scoring procedures is that qualitative data are converted into quantitative units (e.g., word frequencies as in the LIWC software), allowing quantitative analyses. For example, correlating word usage as derived from the LIWC with self- and other ratings in questionnaires allows one to learn more about which linguistic cues might relate to personality traits (e.g., Proyer and Brauer). In the field of personnel selection, a recent study showed that the information derived from written

applications by the LIWC predicted the success of a candidate's application (Brandt and Herzberg).

Furthermore, one might differentiate the types of interviews with regard to their usage and aims in different fields. We will give examples from two disciplines, namely clinical and occupational psychology. In clinical psychology, interviews are often used when the aim is to gather personal information such as acute, past, and process-related information that might be informative for psychological disorders and symptoms. For example, the client's description of life events gives the interviewer insight into not only overt information about critical events but also how the client experienced and experiences them, their attitudes and feelings, and their behavioral reactions when confronted with sensitive topics. In clinical psychology, two types of interviews are broadly distinguished: (1) the *first interview*, which aims to derive a first impression of the client's matter of concern in order to acquire a working hypothesis for their diagnosis and to plan the therapeutic work and interventions and (2) the *diagnostic interview*, which aims to collect information to refine working hypotheses about potential diagnoses.

Illustration—The First Interview and the Diagnostic Interview in Clinical Psychology

The *first interview* in the context of psychotherapy is typically on the low end of the unstructured to structured dimension in order to give the client the opportunity and space to report all information about their current situation and problems, as well as their relevant psychosocial background. However, it is the interviewer's task to ensure that information relevant to the therapy and the therapist's understanding of the problems is given, which is achieved by specific types of questioning (see, for example, Morrison). A first interview is framed by time restrictions (typically 45 minutes). Another principle of the first interview is to establish a trusting interviewer-client relationship to create an atmosphere in which the client can open up and speak about their problems. It must be noted that the interpersonal relationship is of high importance, as such interviews touch the intimate and personal sphere of clients—supportive interpersonal dynamics (see establishing rapport) can therefore contribute to diagnostic and therapeutic outcomes (e.g., Lambert and Barley; Norcross). In the first interview, the diagnostician gathers a broad range of information on various topics (e.g., family and friends of the client or experiences in childhood) and can adjust the

questions to the client's responses. In comparison, information obtained from a questionnaire is limited to the pre-defined answers (e.g., on a rating scale) for each client.

In contrast to the first interview, the *diagnostic interview* relies on higher levels of structuredness. Using standardized interviews such as the *Structured Clinical Interview for the DSM* (SCID; see First), psychological disorders and their severity can be diagnosed by trained therapists (e.g., Widiger; Williams). In this context, it is crucial to phrase the questions in the same way for every patient because even slight differences in the wording may cause different answers, as discussed in the section on the objectivity, reliability, and validity of interviews. Every person with the same degree of depressive symptoms should be diagnosed with the same severity in a depressive disorder by different interviewers. This objectivity in conducting the interview and assessing the responses is especially important in ensuring a reliable and valid diagnosis in order to provide patients with appropriate treatment.

Illustration—The Structured Job Interview in Personnel Selection

In occupational psychology, the interview is often used as a selection method in the application process to fill open positions with the best suited candidates. Schuler et al. showed in their meta-analysis that interviews are the most-used selection instrument, with analyses of application documents ranking second (99% of the studied companies used these). Moreover, the structured interview (73%) was more frequently used than the unstructured interview (42%). Recently, Schuler's work group again surveyed the application criteria among 318 German organizations in 2017 and 2018 in order to re-evaluate the usage of selection criteria (Armoneit et al.). Interestingly, the structured interview remains an important instrument (73%), whereas the usage of unstructured interviews has decreased (34%) over the past decade. This example from occupational psychology signifies the critical role of the interview in the applied disciplines of psychology. More examples can be found in the fields of, for example, educational and health psychology (see e.g., Morrison; Renner and Jacob).

As discussed previously, structured interviews are associated with higher reliability and validity than unstructured interviews; Armoneit et al.'s findings show that this is reflected in their usage in personnel selection. In this section, we highlight on which basis interviews can be structured using job interviews as

an example. First, before conducting the job interview, it is important to collect information concerning the characteristics of the work and conditions the candidate would work in (i.e., the so-called *job analysis*; Campion et al.). This ensures that the requirements for a successful applicant are known and the most suitable person can be selected. To ensure the best fit between the job position and the applicant, the following aspects should be considered and included in the selection interview (Schuler). Like most interviews, the selection interview begins with an introduction of the attendees and an overview of the course, topics, and duration of the interview. To obtain an initial impression of the applicant, they are asked to present themselves; for example, regarding their educational background, professional career, and prior job experiences, and consecutive questions can be asked. After this self-introduction, the candidate should be encouraged by the interviewer to talk about their professional interests, why they applied to the company, and describe their interest in the open position. In the next step, questions about the interviewees' biography and job experiences are asked, which should be based on the requirements of the position (see job analysis) to examine the candidate's fit in relation to the open post. It is important that the interviewer gives a realistic insight into the job position and company and mentions positive as well as negative aspects. Furthermore, situational questions (e.g., "Imagine your work group has a conflict over unjust distributions of workload. Please tell us about your course of action in such a situation to solve the conflict.") should be used as an indicator of the candidate's potential future behaviors in critical and challenging job situations. This type of question enhances validity because it is specifically job-related and the answers of the candidate are valid predictors of future behavior (e.g., Campion et al.). At the conclusion of the interview, open questions from all attendants can be addressed, and the interviewer should give more information on the further procedure and organization of the selection process.

Good Interviewers

After highlighting the differences concerning the structuredness of interviews and formal factors that affect their objectivity, reliability, and validity, it must be noted that *interviewers* themselves might be considered a method factor, as they differ in their "ability" to conduct interviews. The competencies of the interviewer in conducting a good and comprehensive interview are crucial to the quality of the derived information. Morrison argues that a "good interviewer" should have three main aims: (1) to gather the maximum amount of information possible, that

(2) should be accurate and relevant for the specific context, (3) in the shortest amount of time possible. While following those prime aims of collecting information, the interviewer should establish and maintain a good relationship with the client or patient to create a good working atmosphere (*rapport*). Further, a good interviewer has different perspectives on the situation and uses varying sources of information. For example, an interviewer should be sensitive to behavioral observations as an additional source of information: one might think of an interviewee who explicitly denies feeling nervous while simultaneously showing non-verbal signs of nervousness (e.g., trembling movements). This contrast should be taken into account by the interviewer as it provides incremental behavioral information that is contrary to explicit responses by the interviewee (e.g., Nordgaard et al.). Overall, interviewers are faced with high cognitive demand, as they have to collect information by being receptive to various sources of information and their (in)congruence (e.g., concerning speech, facial expressions, and body gestures) while also establishing *rapport* (e.g., Morrison).

Furthermore, the interviewer should embed the derived information into the context of the interviewee. For example, in clinical interviews, important aspects may be the behavior of the client and its dynamics, social milieu, and biological aspects (e.g., comorbid diagnoses), as well as interactions between such factors. Morrison argues that a good interviewer is prepared to handle different “personalities” in the interview situation (see objectivity and unique dyadic interactions). Accordingly, the level of language used might be adjusted to the interviewee’s verbal abilities (Morrison). Hence, a good interviewer knows *how* to collect the required information for a reliable and valid assessment.

As mentioned previously, the interview process is denoted by subjective impressions for both the interviewer and the interviewee. To derive accurate information, the interviewer should monitor and control their *own* potential cognitive and affective biases. For example, common perception biases are the *halo effect* (i.e., the interviewer focuses strongly on one positive aspect while ignoring other potentially relevant information), *anchoring* (i.e., the interviewer uses prior expectations as an anchor that guides the interview), *confirmation bias* (i.e., first assumptions about a person are sought to be confirmed during the interview by seeking information that validates these beliefs while ignoring contrary evidence), and the *affective heuristic* (i.e., judgments are based on personal preferences or salient aspects such as ethnicity, gender, or social background). To address such potential biases, interviewers often participate in training and supervision sessions which contribute to the familiarity with the situations, potential biases, and different types of interviewees that may cue the interviewer’s biases and/or attitudes (e.g., Bensing and Sluijs; Ventura et al.).

Contrasting the Interview with the Use of Questionnaires

Interviews and self-report questionnaires are valuable assessment methods in psychology. In this section, we compare both methods and show their merits and limitations. In contrast to self-report questionnaires, interviews are more individual-centered, personal, subjective, and flexible because questions can be varied from case to case and tailored to the respondent depending on their circumstances and situation. Most people prefer to talk to a professional than check boxes in an anonymous questionnaire, as they perceive interviews as more pleasant and emotionally rewarding (e.g., Neuschwander et al.). The interviewer can also explain certain questions or ask for a more elaborate answer if the interviewee does not understand the question or gives brief or ambiguous responses. However, initial findings show that questionnaire and interview data typically lead to the same conclusions (e.g., Fairburn and Beglin). Despite these advantages and similarities, interviews are usually more time-consuming and less objective because, as discussed previously, the interviewer's bias or low structuredness may influence the answers of the respondent and their interpretation.

Furthermore, self-report questionnaires can be examined and revised with regard to their objectivity, reliability, and validity on the basis of psychometric analyses and considerations. They might have an advantage when intimate and/or embarrassing topics (e.g., eating behavior in patients with eating disorders) or topics that relate to socially desirable responses (e.g., political attitudes and personal values) are discussed: In questionnaires, respondents can answer anonymously without face-to-face interaction with the interviewer. Additionally, socially desirable responding (i.e., answering in normative ways) might be reduced in questionnaires due to anonymity. However, social desirability is also an issue in questionnaires (e.g., Paulhus), and in some cases the interview might be the assessment method that overcomes this bias, as trained interviewers can create an atmosphere of security and intimacy that allows interviewees to respond truthfully. Contrary to interviews, questionnaires typically limit responses to predefined categories (e.g., a scale might give seven response options ranging from 1 = *do not agree* to 7 = *strongly agree*) that cannot be elaborated upon by the respondent, which can limit the comprehensive collection of information (see Stewart and Newton). Thus, questionnaires are more suitable when questions are simple and clear, and no complex attitudes or behaviors are involved. Also, interviewer errors and biases are excluded in data collection using questionnaires, but this also poses the limitation that no behavioral observations can be made.

In terms of resources, questionnaires are more easily administered in group settings and online, both of which are highly effective standardized ways of collecting large sets of data.

As previously discussed, it is possible to “translate” qualitative information from interviews by quantifying and categorizing information of interest into quantitative units. To do so, participants’ responses need to be coded and categorized before they can be analyzed and interpreted. This is typically done in specific software packages like MAXQDA, which allows researchers to code and analyze interview data based on predefined categories of interest (e.g., the occurrence of certain events) and using computerized language analyses (e.g., LIWC; Tausczik and Pennebaker).

Although interviews and questionnaires share similarities and differ with regard to their merits and disadvantages, it is questionable whether one method is superior to the other. We argue that this question relies on the aims of data collection, resources, and type of hypothesis (i.e., exploratory vs. confirmatory). Again, findings on the comparison of questionnaire and interview methods to assess the same variable of interest have shown that both converge comparatively well (e.g., Fairburn and Beglin; Widiger).

Conclusion

Our discussion has shown that the interview method is a valuable approach to collecting diagnostic information in research and applied psychology. As meta-analyses have demonstrated, interviews allow for comparatively reliable assessments and provide valid and oftentimes incremental information, even beyond objective criteria (Schmidt and Hunter). The knowledge of the objectivity, reliability, and validity of interviews as well as factors systematically affecting them has advanced the understanding of the shortcomings of interviews and facilitated the adjustment of interview strategies; for example, by maximizing the structuredness, using multiple observers, using recordings of interviews, and increasing efforts to train interviewers in conducting them and processing the collected information (e.g., using standardized coding rules).

Putting the merits and shortcomings of interviews and other assessment methods aside, we want to highlight that the assessment of psychological phenomena, clinical diagnoses, or personnel selection should never rely or be based on a single method. Numerous studies have demonstrated that the use of multiple assessments with a variety of methods provides the best approach to collecting information and deriving reliable and valid conclusions (e.g., Campbell and

Fiske; Fairburn and Beglin; Schmidt and Hunter; Vazire; Widiger). For example, by supplementing interviews with questionnaire data, information provided by the interviewee's knowledgeable others (e.g., teachers or supervisors), and objective data (e.g., [neuro]psychophysiological measures). Taking the many merits of interviews into account, we conclude that the interview is a strong method in the field of psychological assessment.

Works Cited

- Armoneit, Corinna, et al. "Nutzung, Validität, Praktikabilität und Akzeptanz psychologischer Personalauswahlverfahren in Deutschland 1985, 1993, 2007, 2020: Fortführung einer Trendstudie." [Use, Validity, Practicality, and Acceptance of Personnel Selection Methods in Germany 1985, 1993, 2007, 2020: The Continuation of a Trend Study] *Zeitschrift für Arbeits- und Organisationspsychologie*, vol. 64, no. 2, 2020, pp. 67–82, <http://dx.doi.org/10.1026/0932-4089/a000311>.
- Bensing, Jozen M., and Emmy M. Sluijs. "Evaluation of an Interview Training Course for General Practitioners." *Social Science & Medicine*, vol. 20, no. 7, 1985, pp. 737–44, [https://doi.org/10.1016/0277-9536\(85\)90064-4](https://doi.org/10.1016/0277-9536(85)90064-4).
- Brandt, Pia M., and Philipp Y. Herzberg. "Is a Cover Letter Still Needed? Using LIWC to Predict Application Success." *International Journal of Selection and Assessment*, vol. 28, no. 4, 2020, pp. 417–29, <https://doi.org/10.1111/ijsa.12299>.
- Brauer, Kay, and René T. Proyer. "Dyadic Processes." *Encyclopedia of Evolutionary Psychological Science*, edited by Todd Shackelford and Viviana A. Shackelford, Springer, 2021, pp. 2167–74, http://dx.doi.org/10.1007/978-3-319-16999-6_1720-1.
- Bund Deutscher Psychologen. *Berufsbild Psychologie: Psychologische Tätigkeitsfelder*. [Job Description Psychology: Psychological Fields of Activity], Berufsverband Deutscher Psychologinnen und Psychologen (BDP) e.V., 2017, www.bdp-verband.de/fileadmin/user_upload/BDP/verband/Untergliederungen/Sektionen/AusFort-und>Weiterbildung_in_Psychologie/PDF/bdp-broschure-berufsbild-psychologie.pdf.
- Burnett, Jennifer R., et al. "Interview Notes and Validity." *Personnel Psychology*, vol. 51, no. 2, 1998, pp. 375–96, <https://doi.org/10.1111/j.1744-6570.1998.tb00730.x>.
- Bollmann, Stella. "Exploratory Factor Analysis." *Encyclopedia of Personality and Individual Differences*, edited by Virgil Zeigler-Hill and Todd Shackelford, Springer, 2020, pp. 1497–502, https://doi.org/10.1007/978-3-319-28099-8_1302-2.
- Campbell, Donald T., and Donald W. Fiske. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin*, vol. 56, no. 2, 1959, pp. 81–105, <https://psycnet.apa.org/doi/10.1037/h0046016>.
- Campion, Michael A., et al. "A Review of Structure in the Selection Interview." *Personnel Psychology*, vol. 50, no. 3, 1997, pp. 655–702, <https://doi.org/10.1111/j.1744-6570.1997.tb00709.x>.
- Conway, James, et al. "A Meta-Analysis of Interrater and Internal Consistency Reliability of Selection Interviews." *Journal of Applied Psychology*, vol. 80, no. 5, 1995, pp. 565–79, <https://psycnet.apa.org/doi/10.1037/0021-9010.80.5.565>.

- Costa, Paul T. Jr., and Robert R. McCrae. "The Revised NEO Personality Inventory (NEO-PI-R)." *The SAGE Handbook of Personality Theory and Assessment: Personality Measurement and Testing*, edited by Gregory J. Boyle et al., vol. 2, SAGE, 2008, pp. 179–98, <https://doi.org/10.4135/9781849200479>.
- Craig, Robert J. *Clinical and Diagnostic Interviewing*. Aronson, 2005.
- Fairburn, Christopher G., and Sarah J. Beglin. "Assessment of Eating Disorders: Interview or Self-Report Questionnaire?" *International Journal of Eating Disorders*, vol. 16, no. 4, 1994, pp. 363–70, [https://doi.org/10.1002/1098-108X\(199412\)16:4%3C363::AID-EAT2260160405%3E3.0.CO;2-%23](https://doi.org/10.1002/1098-108X(199412)16:4%3C363::AID-EAT2260160405%3E3.0.CO;2-%23).
- First, Michael B. "Structured Clinical Interview for the DSM (SCID)." *The Encyclopedia of Clinical Psychology*, Wiley, 2015, *Wiley Online Library*, <https://doi.org/10.1002/9781118625392.wbecp351>.
- Fisher, Seymour, and Roger P. Greenberg, *Freud Scientifically Reappraised: Testing the Theories and Therapy*. Wiley, 1996.
- Furr, Michael R., and Verne R. Bacharach. *Psychometrics: An Introduction*. SAGE, 2014.
- Harari, Gabriella M., et al. "Smartphone Sensing Methods for Studying Behavior in Everyday Life." *Current Opinion in Behavioral Sciences*, vol. 18, 2017, pp. 83–90, <https://doi.org/10.1016/j.cobeha.2017.07.018>.
- Henson, Ramon, et al. "Effects of Interviewer Style on Quality of Reporting in a Survey Interview." *The Journal of Psychology: Interdisciplinary and Applied*, vol. 92, no. 2, 1976, pp. 221–27, <https://doi.org/10.1080/00223980.1976.9915816>.
- Huffcutt, Allen I., et al. "Employment Interview Reliability: New Meta-Analytic Estimates by Structure and Format." *International Journal of Selection and Assessment*, vol. 21, no. 3, 2013, pp. 264–76, <https://doi.org/10.1111/ijsa.12036>.
- Jiménez, Paulino, and Erich Raab. "Das Berufsbild von Psychologen im Vergleich mit anderen Berufen." [The Job Profile of Psychologists in Comparison with Other Professions], *Psychologische Rundschau*, vol. 50, no. 1, 1999, pp. 26–32, <https://doi.org/10.1026//0033-3042.50.1.26>.
- Lalande, Kathleen M., and Bonanno, George A. "Retrospective Memory Bias for the Frequency of Potentially Traumatic Events: A Prospective Study." *Psychological Trauma: Theory, Research, Practice, and Policy*, vol. 3, no. 2, 2011, pp. 165–70, <https://psycnet.apa.org/doi/10.1037/a0020847>.
- Latham, Gary P., and Lise M. Saari. "Do People Do What They Say? Further Studies on the Situational Interview." *Journal of Applied Psychology*, vol. 69, no. 4, 1984, pp. 569–73, <https://psycnet.apa.org/doi/10.1037/0021-9010.69.4.569>.
- Lambert, Michael J., and Dean E. Barley. "Research Summary on the Therapeutic Relationship and Psychotherapy Outcome." *Psychotherapy: Theory, Research, Practice, Training*, vol. 38, no. 4, 2011, pp. 357–61, <https://psycnet.apa.org/doi/10.1037/0033-3204.38.4.357>.
- Loftus, Elizabeth F., and John, C. Palmer. "Reconstruction of Automobile Destruction: An Example of the Interaction between Language and Memory." *Journal of Verbal Learning and Verbal Behavior*, vol. 13, no. 5, 1974, pp. 585–89, [https://psycnet.apa.org/doi/10.1016/S0022-5371\(74\)80011-3](https://psycnet.apa.org/doi/10.1016/S0022-5371(74)80011-3).
- Mandler, George. *A History of Modern Experimental Psychology: From James and Wundt to Cognitive Science*. MIT Press, 2011.
- McAdams, Dan P. "Narrative Identity: What Is It? What Does It Do? How Do You Measure It?" *Imagination, Cognition and Personality*, vol. 37, no. 3, 2018, pp. 359–72, <https://doi.org/10.1177/0276236618756704>.

- McAdams, Dan P., et al. *Identity and Story: Creating Self in Narrative*. American Psychological Association, 2006.
- McCrae, Robert R., and Paul T. Jr. Costa. "Validation of the Five-Factor Model of Personality Across Instruments and Observers." *Journal of Personality and Social Psychology*, vol. 52, no. 1, 1987, pp. 81–90, <https://doi.org/10.1037//0022-3514.52.1.81>.
- Michell, John. *An Introduction to the Logic of Psychological Measurement*. Psychological Press, 2014.
- Miller, Paul R., et al. "Inpatient Diagnostic Assessments: 1. Accuracy of Structured vs. Unstructured Interviews." *Psychiatry Research*, vol. 105, no. 3, 2001, pp. 255–64, [https://doi.org/10.1016/s0165-1781\(01\)00317-1](https://doi.org/10.1016/s0165-1781(01)00317-1).
- Mischel, Theodore. "Wundt and the Conceptual Foundations of Psychology." *Philosophy and Phenomenological Research*, vol. 31, no. 1, 1970, pp. 1–26, <https://psycnet.apa.org/doi/10.2307/2105977>.
- Morrison, James R. *The First Interview*. Guilford, 2014.
- Neuschwander, Murielle, et al. "Acceptance of a Structured Diagnostic Interview in Children, Parents, and Interviewers." *International Journal of Methods in Psychiatric Research*, vol. 26, no. 3, 2017, e1573, <https://doi.org/10.1002/mp.1573>.
- Norcross, John C. "The Therapeutic Relationship." *The Heart and Soul of Change: Delivering What Works in Therapy*, edited by Barry L. Duncan et al., American Psychological Association, 2010, pp. 113–41, <https://psycnet.apa.org/doi/10.1037/12075-004>.
- Nordgaard, Julie, et al. "The Psychiatric Interview: Validity, Structure, and Subjectivity." *European Archives of Psychiatry and Clinical Neuroscience*, vol. 263, 2013, pp. 353–64, <https://doi.org/10.1007/s00406-012-0366-z>.
- Paulhus, Delroy L. "Measurement and Control of Response Bias." *Measures of Personality and Social Psychological Attitudes*, edited by John P. Robinson et al., pp. 17–59. *Measures of Social Psychological Attitudes*, vol. 1, Academic Press, 1999. <https://psycnet.apa.org/doi/10.1016/B978-0-12-590241-0.50006-X>.
- Platt, Tracey, et al. "Gelotophobia in Practice and Implications of Ignoring it." *European Journal of Humour Research*, vol. 4, no. 2, 2016, pp. 46–56, <https://doi.org/10.7592/EJHR2016.4.2.platt>.
- Powell, Martine B., et al. "Skill in Interviewing Reduces Confirmation Bias." *Journal of Investigative Psychology and Offender Profiling*, vol. 9, no. 2, 2014, pp. 126–34, <https://doi.org/10.1002/jip.1357>.
- Proyer, René T., and Kay Brauer. "Exploring Adult Playfulness: Examining the Accuracy of Personality Judgments at Zero-Acquaintance and an LIWC Analysis of Textual Information." *Journal of Research in Personality*, vol. 73, 2018, pp. 12–20, <https://doi.org/10.1016/j.jrp.2017.10.002>.
- Read, J. Don, and Deborah A. Connolly. "The Effects of Delay on Long-Term Memory for Witnessed Events." *Memory for Events*, edited by Michael P. Toglia, et al., pp. 117–55. *The Handbook of Eyewitness Psychology*, vol. 1, Lawrence Erlbaum, 2007.
- Renner, Karl-Heinz, and Nora-Corinna Jacob. *Das Interview: Grundlagen und Anwendung in Psychologie und Sozialwissenschaften* [The Interview: Basics and Application in Psychology and Social Sciences], Springer, 2020, <https://doi.org/10.1007/978-3-662-60441-0>.
- Rogers, Richard. *Diagnostic and Structured Interviewing: A Handbook for Clinical Practice*. Guilford, 2001.

- Rogers, Richard. "Standardizing DSM-IV Diagnoses: The Clinical Applications of Structured Interviews." *Journal of Personality Assessment*, vol. 81, no. 3, 2003, pp. 250–55, https://doi.org/10.1207/S15327752JPA8103_04.
- Ruch, Willibald, and René T. Proyer. "The Fear of Being Laughed At: Individual and Group Differences in Gelotophobia." *Humor: International Journal of Humor Research*, vol. 21, no. 1, 2008, pp. 47–67, <https://doi.org/10.1515/HUMOR.2008.002>.
- Schmidt, Frank L., and John E. Hunter. "The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings." *Psychological Bulletin*, vol. 124, no. 2, 1998, pp. 262–74, <http://dx.doi.org/10.1037/0033-2909.124.2.262>.
- Schuler, Heinz. *Das Einstellungsinterview*. [The Selection Interview], Hogrefe, 2018.
- Schuler, Heinz, et al. "Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen" [The Use of Personnel Selection Instruments in German Organizations in the Last 20 Years], *Zeitschrift für Personalpsychologie*, vol. 6, no. 2, 2007, pp. 60–70, <https://doi.org/10.1026/1617-6391.6.2.60>.
- Smyth, Joshua M., and Arthur A. Stone. "Ecological Momentary Assessment in Behavioral Medicine." *Journal of Happiness Studies*, vol. 4, 2003, pp. 35–52, <https://doi.org/10.1023/A:1023657221954>.
- Stewart, Abigail J., and Nicky J. Newton. "Devilish Details: The Importance of Marginalia in Personality Research." *Journal of Research in Personality*, vol. 97, 2022, e104189, <https://doi.org/10.1016/j.jrp.2022.104189>.
- Stocco, Corey S., et al. "Improving the Interview Skills of College Students Using Behavioral Skills Training." *Journal of Applied Behavior Analysis*, vol. 50, no. 3, 2017, pp. 495–510, <https://doi.org/10.1002/jaba.385>.
- Sun, Hanyu, et al. "The Relationship between Interviewer-Respondent Rapport and Data Quality." *Journal of Survey Statistics and Methodology*, vol. 9, no. 3, 2021, pp. 429–48, <https://doi.org/10.1093/jssam/snz043>.
- Tausczik, Yla R., and James W. Pennebaker. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology*, vol. 29, no. 1, 2010, pp. 24–54, <http://dx.doi.org/10.1177/0261927X09351676>.
- Tinsely, Howard E. A., and David J. Weiss. "Interrater Reliability and Agreement of Subjective Judgments." *Journal of Counseling Psychology*, vol. 22, no. 4, 1975, pp. 358–76, <https://psycnet.apa.org/doi/10.1037/h0076640>.
- Vazire, Simine. "Who Knows What about a Person? The Self-Other Knowledge Asymmetry (SOKA) Model." *Journal of Personality and Social Psychology*, vol. 98, no. 2, 2010, pp. 281–300, <https://doi.org/10.1037/a0017908>.
- Ventura, Joseph, et al. "Training and Quality Assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P)." *Psychiatry Research*, vol. 79, no. 2, 1998, pp. 163–73, [https://doi.org/10.1016/S0165-1781\(98\)00038-9](https://doi.org/10.1016/S0165-1781(98)00038-9).
- Vincent, Douglas F. "The Origin and Development of Factor Analysis." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 2, no. 2, 1953, pp. 107–17.
- Walker, David A. "Estimating How Many Observations Are Needed to Obtain a Required Level of Reliability." *Journal of Modern Applied Statistical Methods*, vol. 7, no. 1, 2008, pp. 152–57, <http://dx.doi.org/10.22237/jmasm/1209615060>.
- Widiger, Thomas A. "Clinical Interviews." *Evidence-Based Outcome Research: A Practical Guide to Conducting Randomized Controlled Trials for Psychosocial Interventions*, edited by Arthur M. Nezu and Christine M. Nezu, Oxford UP, 2008, pp. 47–66.

- Williams, Janet B. W. "A Structured Interview Guide for the Hamilton Depression Rating Scale." *Archives of General Psychiatry*, vol. 45, no. 8, 1988, pp. 742–47, <https://doi.org/10.1001/archpsyc.1988.01800320058007>.
- Wittchen, Hans-Ulrich. "Reliability and Validity Studies of the WHO-Composite International Diagnostic Interview (CIDI): A Critical Review." *Journal of Psychiatric Research*, vol. 28, no. 1, 1994, pp. 57–84, [https://doi.org/10.1016/0022-3956\(94\)90036-1](https://doi.org/10.1016/0022-3956(94)90036-1).
- Wu, Kevin D., and Lee A. Clark. "Relations between Personality Traits and Self-Reports of Daily Behavior." *Journal of Research in Personality*, vol. 37, no. 4, 2003, pp. 231–56, [https://doi.org/10.1016/S0092-6566\(02\)00539-1](https://doi.org/10.1016/S0092-6566(02)00539-1).
- Young, Jacy L. "Thinking in Multitudes: Questionnaires and Composite Cases in Early American Psychology." *History of the Human Sciences*, vol. 33, no. 3–4, 2020, pp. 160–74, <https://doi.org/10.1177/0952695120903909>.
- Zimbardo, Phillip G., et al. *Psychology: Core Concepts*. Pearson, 2014.