American Journal of Clinical Hypnosis

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ujhy20

Transcultural Factors in Hypnotizability Scales: Limits and Prospects

Claire M. Champigny & Amir Raz

a McGill University, Montréal, Canada

Published online: 11 Aug 2015.

To cite this article: Claire M. Champigny & Amir Raz (2015) Transcultural Factors in Hypnotizability Scales: Limits and Prospects, American Journal of Clinical Hypnosis, 58:2, 171-194, DOI: 10.1080/00029157.2015.1061473

To link to this article: http://dx.doi.org/10.1080/00029157.2015.1061473

PLEASE SCROLL DOWN FOR ARTICLE
Transcultural Factors in Hypnotizability Scales: Limits and Prospects

Claire M. Champigny and Amir Raz
McGill University, Montréal, Canada

Hypnotic suggestibility—loosely termed hypnotizability—is difficult to assess across cultures. Investigators often use translated research instruments to guide their inquiry in disparate geographic locations. Present-day hypnosis researchers rely heavily on two primary scales that are more than half a century old: the Stanford Hypnotic Susceptibility Scale: Form C (SHSS:C) (Weitzenhoffer & Hilgard, 1959) and the Harvard Group Scale of Hypnotic Susceptibility: Form A (HGSCHS:A) (Shor & Orne, 1962). Scholars typically translate these scales to measure hypnotizability transculturally. This approach, however, operates under the specious assumption that the concept of hypnotizability is largely monolithic or universal across cultures. Whereas translations likely conserve the linguistic content, they may arguably imply different cultural meanings and historical subtexts. Whereas social scientists acknowledge the importance of qualitative and phenomenological accounts in the study of altered consciousness, including suggestibility, researchers interested in hypnotizability consider the impact of findings from anthropology and ethnography too little. Clinicians and scholars of hypnosis would stand to benefit from incorporating the insights afforded by transcultural research in the overarching investigation of a concept as nuanced as hypnotizability.

Keywords: ethnography, hypnotizability, transcultural research, translation

Hypnosis and Hypnotizability

Hypnosis, including susceptibility to suggestion, compliance with instructions, and adherence to directives, has steadily harnessed scientific attention over the past century (see Figure 1) (Baker, 1990; Brown & Fromm, 1986; Halligan & Oakley, 2014; Hull, 1933; Kihlstrom, 2013; Kirsch, 2014; Raz, 2007; Wagstaff, 1981). Whereas individuals likely vary in their experience of hypnosis, they typically report an increase in absorption, focused attention, and reduction in spontaneous thoughts (Oakley & Halligan, 2009). As such, a hypnotic experience may foster powerful self-regulation, akin to other meditative practices (Lifshitz, Cusumano, & Raz, 2014; Raz & Lifshitz, in press; Raz, Shapiro, Fan, &Posner, 2002). Although hypnosis has been used for centuries, it has only been relatively recently that research has provided empirical evidence...
FIGURE 1 Scopus publication profile for “hypnosis” for “all fields” (article and review) since 1900.

showing how hypnosis can help in gaining control over deeply-ingrained psychological processes (Raz, 2011). For example, cognitive accounts demonstrate that hypnotic and post-hypnotic suggestions emitted during a hypnotic experience can: (1) disrupt word recognition (e.g., Lifshitz, Aubert-Bonn, Fischer, Kashem, & Raz, 2012), (2) generate or alter the symptomatology of neuropsychological conditions (e.g., Barabasz & Barabasz, 2008; Cohen-Kadosh, Henik, Catena, Walsh, & Fuentes, 2009; Halligan & Oakley, 2013; Kihlstrom, 2013; Oakley & Halligan, 2009, 2013; Raz, 2004; Terhune, Cardeña, & Lindgren, 2010), and (3) override highly automatic multimodal sensory integration (e.g., Déry, Campbell, Lifshitz, & Raz, 2014). Research suggests that distinct patterns of brain activations attributable to hypnosis affect cognitive processing, perception, and ideomotor action (Del Casale et al., 2012; Egner, Jamieson, & Gruzelier, 2005; Fingelkurts, Fingelkurts, Kallio, & Revonsuo, 2007; Halligan & Oakley, 2013; Jamieson, 2007; Kihlstrom, 2013; Landry & Raz, 2015; Mazzoni, Venneri, McGeown, & Kirsch, 2013; Oakley & Halligan, 2009, 2013; Rainville, Hofbauer, Bushnell, Duncan, & Price, 2002).

Individual variability in hypnotic response often goes by the appellation hypnotizability (Carhart-Harris et al., 2015; Halligan & Oakley, 2014; Laurence, Beaulieu-Prévost, & Du Chéné, 2008; Raz, 2007). The American Psychological Association Division 30 defines hypnotizability as “an individual’s ability to experience suggested alterations in physiology, sensations, emotions, thoughts, or behavior during hypnosis” (Elkins, Barabasz, Council, & Spiegel, 2015, p. 6). In this article the authors center on this susceptibility to hypnotic suggestion and common caveats associated with its measurement and assessment across cultures.
Measuring Hypnotizability

Individual differences in hypnotic response are prevalent. “Most of those who have shaped the history of hypnosis not only recognize the major importance of these individual differences in responsivity to suggestions, but also attempted to document them” (Dixon & Laurence, 1992, p. 36). The scientific measurement and quantification of such parameters is often elusive and tenuous (Woody & Barnier, 2008). The disposition toward a hypnotic response—hypnotizability—provides a starting point from which to quantify hypnotic phenomena under standard conditions. While some hypnosis researchers disagree on the nuances of the term hypnotizability (Barabasz & Perez, 2007; Council, 1999; Kirsch, 2014; Kirsch et al., 2011), here we conceptualize the term as responsivity to hypnotic suggestion.

In order to measure such phenomena, and in accordance with traditional scientific models, researchers devised scales that assign numbers to the varying manifestations of hypnotizability (e.g., Barry, MacKinnon, & Murray, 1931; Davis & Husband, 1931; Friedlander & Sarbin, 1938; White, 1930). Although several subjective scales for assessing hypnotizability exist (e.g., the Subjective Experiences Scale (Kirsch, Council, & Wickless, 1990) and the Inventory Scale of Hypnotic Depth (Field, 1965), researchers in the United States have come to rely on two hypnotizability scales, which together have become the gold standard for assessing hypnotizability in research (Register & Kihlstrom, 1986). In this article, therefore, we focus on these two prominent instruments. Psychologists André Weitzenhoffer and Ernest Hilgard developed the Stanford Hypnotic Susceptibility Scale: Form C (SHSS:C) in 1959, whereas psychologists Ronald Shor and Emily Orne created the Harvard Group Scale of Hypnotic Susceptibility: Form A (HGSHS:A) in 1962. On the one hand, the SHSS:C evaluates individual subjects on 12 hypnotic suggestions. Throughout the hypnotic procedure, the hypnotist takes notes to validate subject responses. The SHSS:C displays particularly impressive test–retest reliability over 10 ($r = .64$), 15 ($r = .82$), and 25 years ($r = .71$) (Piccione, Hilgard, & Zimbardo, 1989). Researchers consider the SHSS:C a reliable and robust measure of individual responsiveness. On the other hand, the HGSHS:A tests multiple subjects simultaneously using 12 hypnotic suggestions and a self-scoring questionnaire that subjects complete immediately following their hypnotic experience. This scale provides a method for quickly identifying high and low hypnotizable individuals within a large sample, although it rates individuals with lower sensitivity than does the SHSS:C. Several experts, most notably Weitzenhoffer himself (1974, 1978, 1980), have expressed their concerns regarding the confusion between suggestions and instructions, as well as the clinical impracticality (50–90 minutes) of the SHSS:C. Despite the development of a plethora of other scales (see Table 1), the SHSS:C and the HGSHS:A prevail, especially when used in concert, as the most functional and psychometrically robust hypnotizability scales (Barber, 1965; Hilgard, 1965; Sheehan & McConkey, 1982; Spanos & Chaves, 1991).
TABLE 1
Hypnotizability and Suggestibility Scales in Use Today

<table>
<thead>
<tr>
<th>Scale</th>
<th>Country of origin</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford Hypnotic Susceptibility Scale: Form C (SHSS: C)</td>
<td>USA</td>
<td>Gold standard for individual hypnotizability measure.</td>
</tr>
<tr>
<td>(Weitzenhoffer &amp; Hilgard, 1962)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harvard Group Scale of Hypnotic Susceptibility: Form A (HGSS: A)</td>
<td>USA</td>
<td>Gold standard for group hypnotizability measure.</td>
</tr>
<tr>
<td>(Shor &amp; Orne, 1962)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford Hypnotic Susceptibility Scale: Forms A and B (SHSS: A and B)</td>
<td>USA</td>
<td>The original version of Form C. These scales are essentially equivalent, with only slight changes in wording and procedure.</td>
</tr>
<tr>
<td>(Weitzenhoffer &amp; Hilgard, 1959)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford Hypnotic Clinical Scale (SHCS)</td>
<td>USA</td>
<td>Brief scale suitable for individuals with limited movement. Designed to maximize the probability of a successful hypnotic experience and to provide useful clinical information.</td>
</tr>
<tr>
<td>(Morgan &amp; Hilgard, 1978)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford Profile Scales of Hypnotic Susceptibility: Forms I and II (SPHS: I and II)</td>
<td>USA</td>
<td>Designed to provide a profile of hypnotic abilities, such as hallucinations and cognitive distortions.</td>
</tr>
<tr>
<td>(Weitzenhoffer &amp; Hilgard, 1963)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revised Stanford Profile Scales of Hypnotic Susceptibility: Forms I and II (R-SPHS: I and II)</td>
<td>USA</td>
<td>Revised SPHS: I and II.</td>
</tr>
<tr>
<td>(Weitzenhoffer &amp; Hilgard, 1967)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford Hypnotic Arm Levitation Induction and Test (SHALIT)</td>
<td>USA</td>
<td>A single item test, consisting of hypnotic induction/arm levitation, designed for clinical use and as a screening measure.</td>
</tr>
<tr>
<td>(Hilgard, Crawford, &amp; Wert, 1979)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creative Imagination Scale (CIS)</td>
<td>USA</td>
<td>Administration with or without hypnotic induction. Focuses on non-hypnotic suggestibility.</td>
</tr>
<tr>
<td>(Barber &amp; Wilson, 1978)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barber Suggestibility Scale (BSS)</td>
<td>USA</td>
<td>Short scale, abandoned in favor of the CURSS.</td>
</tr>
<tr>
<td>(Barber &amp; Wilson, 1978)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypnotic Induction Profile (HIP)</td>
<td>USA</td>
<td>Developed for clinical settings. Includes a brief but controversial “eye roll test.”</td>
</tr>
<tr>
<td>(Spiegel, 1974)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carleton University Responsiveness to Suggestions Scale (CURSS)</td>
<td>Canada</td>
<td>Short scale. Covers ideomotor and cognitive suggestions.</td>
</tr>
<tr>
<td>(Spanos, 1983)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waterloo-Stanford Group C Scale (WSGC)</td>
<td>Canada</td>
<td>Group adaptation of SHSS:C modified for easier group presentation and self-scoring.</td>
</tr>
<tr>
<td>(Bowers, 1993, 1998)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Role of Culture in Altered States of Consciousness

Social cognitive theories of hypnosis assert that response expectancy, which culture largely mediates, plays a predominant role in shaping response to hypnotic suggestions (Barber, Spanos, & Chaves, 1974; Kirsch, 1991; Kirsch & Lynn, 1995, 1997, 1999; Kirsch, Silva, Comey, & Reed, 1995; Kirsch, Wickless, & Moffitt, 1999; Lynn & Green,
Transcultural Factors in Hypnotizability Scales

2011; Lynn, Kirsch, & Hallquist, 2008; Lynn & Rhue, 1991; Page, Handley, & Green, 1997; Sheehan & Perry, 1977; Spanos & Chaves, 1989; Wagstaff, David, Kirsch, & Lynn, 2010). Simply speaking, culture consists of “regular occurrences in the humanly created world, in the schemas people share as a result of these, and in the interactions between these schemas and this world” (Strauss & Quinn, 1997, p. 7). In other words, culture may be viewed as an organized system of knowledge and beliefs that allows a group to structure its experiences (Tseng, 1997). Culture remains a fuzzy concept because anthropologists and other researchers in related fields shy away from consensus on boundaries and labels, such as deciding what types of shared experiences fall within the purview of culture (Strauss & Quinn, 1997). For the purpose of the present discussion, however, we define culture as referring to the unifying beliefs, customs, and arts of a particular group of people.

Culture mediates the ability to imagine a suggested experience, which further depends on the interaction between neurocognitive predispositions and socio-cultural beliefs about altered states of consciousness (Bourguignon & Evascu, 1977; Krippner, 2000). For example, labeling an induction procedure as hypnosis rather than relaxation increases participant response to subsequent suggestions (Hylands-White & Derbyshire, 2007; Oakley & Gandhi, 2005). Certain cultural activities may induce specific altered states of consciousness, such as the use of psychoactive plants, fasting, thirsting, self-mutilation, sweat lodges, sleeplessness, incessant dancing, bleeding, walking on hot coals, meditation, chanting, or drumming (Furst, 1977). These activities resemble hypnotic practices in the sense that altered cognition and personal beliefs co-determine the desired experience. Moreover, such practices prevail globally; approximately 89% of 488 studied societies socially approve and promote altered states of consciousness (Bourguignon & Evascu, 1977). This social encouragement primes certain populations for hypnotic phenomena by ingraining in them a sense of familiarity with such psychological experiences. By promoting altered states of consciousness, populations may shape their own mental capacities and render themselves more hypnotizable.

When comparing hypnotic-like experiences cross-culturally, researchers may inappropriately assume that hypnosis, a concept specific to Western culture, readily translates into a similar construct in other cultures (Cardeña & Krippner, 2010). However, “no matter how many features [two phenomena] share . . . there is no escaping the inevitable fact that at some point there will be a ‘conceptual fork’ in the road, where the phenomena will differ” (Shiraev & Levy, 2013, p. 40). Researchers stand to benefit from a scientific understanding of the sociocultural factors that render certain populations more prone to hypnotic-like experiences, compared to other groups. Because cultural factors affect response expectancy, which in turn helps shape individual hypnotizability, hypnotizability likely varies across societies and cultures; however, evidence to this effect has yet to materialize.
Lost in Translation

Translating quantitative instruments—such as hypnosis scales—constitutes an essential element of any transcultural study (Kleinman, 1987). A frequent error, however, posits that a concept—in this case, hypnotizability—is universal, and therefore measurable with the same test items across cultures by translating the original language into the language of the target population (Jones & Kay, 1992). Such an ethnocentric error demonstrates that Western researchers often fail to assess their own sociocultural background, despite the fact that this consideration would have optimized analyses of the culture under study (Stanghellini & Ciglia, 2013). Although translation procedures aim for rigor and linguistic accuracy, they tend to focus on the literal translation of phrases and often miss the complex subtleties of cultural interpretations concerning higher order abstractions and concepts (e.g., Bravo, Canino, Rubio-Stipec, & Woodbury, 1991; Ketzer & Crescenzi, 2002). The fields of biology and medicine tend to neglect the fact that perceptual, cognitive, interpersonal, and social processes mediate the verbal explanations of bodily processes in an individual (Kirmayer, 2005). For example, cultural explanations for certain somatic feelings set up expectations that influence the ways in which individuals attend to their mind–body continuum (Kirmayer & Sartorius, 2007). Koreans suffering from abdominal pain may feel a burning sensation because they associate such pain with Hwa-Byung, or fire illness, a syndrome based on the notion of an imbalance of fire, a basic constituent of the body (Lin, 1983; Lin et al., 1992). These culturally-based expectations may partially account for the diversity pertaining to somatic and delusional disorders (American Psychiatric Association [APA], 2013). Based on the current translation methodology of hypnotizability scales, it seems that researchers often assume that the description of embodied experience correctly and objectively reflects that experience (cf., Kirmayer, 2003). As a result, researchers often resort to a simple translation when they should instead re-work the test items to convey the appropriate concepts pertaining to each culture (Kleinman, 1987).

Translation Procedures

Researchers have an array of options at their disposal for the translation of a scientific instrument. In back translation, one of the most common translation procedures, one interpreter translates the first half of the instrument from the source language to the target language, while a second interpreter performs the same task on the second half of the instrument (Brislin, 1970; Chapman & Carter, 1979; Jones & Kay, 1992). They then switch positions and translate each text back to the source language, enabling a comparison of the two English versions of the instrument. After discussing the content of the items with the investigator, the interpreters repeat the back translation until they deem both versions equivalent. Over time, translation procedures have changed. In current studies, researchers often ask an interpreter to translate the entire instrument to the target language and then back-translate each section into the source language.
language, then ask another, unbiased interpreter—who should have no knowledge of the study—to translate the new version back to the original language (Wild et al., 2005). Researchers then closely compare the back-translated version with the original version in order to highlight and correct discrepancies over the course of several more back translation procedures.

The World Health Organization (WHO, 2015b) recommends adding an additional step to this translation procedure consisting of pretesting, a process wherein researchers administer the translated tool to individuals representative of the target audience. After their debriefing, the respondents provide feedback about each question: was it clearly stated? How did they understand it? Could they repeat the question in their own words? Did anything specific come to mind when they heard a particular phrase? Regrettably, a large majority of cross-cultural hypnosis studies do not pre-test translated hypnotizability scales. Moreover, despite its meticulousness, back translation ensures only literal accuracy (Allen & Walsh, 2000) and exhibits weak conceptual equivalence (Larkin, Dierckx de Casterlé, & Schotsmans, 2007). Conceptual equivalence addresses the translation of not only words, but also their embedded meaning and intent; this notion emphasizes the importance for a measured construct to uphold the appropriate meaning between the languages under study (Brislin, 1993; Hunt & Bhopal, 2004).

Even when translation procedures barely attain conceptual equivalence in research materials—such as hypnotizability scales—participants will still likely understand the translated scale differently than they would have understood the original scale. A recent examination of the pass rates for the posthypnotic amnesia item in the HGSHS:A across 10 countries has exposed problems in internal consistency reliability, as illustrated by highly fluctuating rates depending on the language of the instrument (see Figure 2) (Freedman, 2012). The posthypnotic amnesia item has the most variable pass rate of

![FIGURE 2 Percentage of pass rate for amnesia item of HGHS:A.](image-url)
any suggestion in the HGSHS:A, ranging from 13% in Israel to 71% in Denmark (Lichtenberg, 2008; Zachariae, Sommerlund, & Molay, 1996). For this suggestion, the hypnotist instructs participants to forget what has happened during the session until the onset of a specific cue. As soon as the hypnotic experience ends, they list “all the things that happened since you began looking at the target” in their response booklet. Participants who report fewer than four hypnotic suggestions have successfully passed the posthypnotic amnesia item. In other words, participants can technically pass this item not because they have temporarily forgotten the suggestions, but because they have failed to report these suggestions (Cooper, 1972).

As a result, researchers at Concordia University investigated the reasons for such unexpected variability in pass rates, and attributed the variations to translational and cultural differences (Freedman, 2012). For example, they conjectured that the unusually high passing rate (53%) in a Finnish population may have arisen due to a lack of conceptual equivalence (Kallio & Ihamuotila, 1999). They reported that “the wording [in the Finnish scale] . . . may lead to a somewhat different meaning when translated. In Finnish the verb happen (tapahuida) has a passive connotation (more like occur) and a better translation might have been the more active tehdä which is equal to do” (Kallio & Ihamuotila, 1999, p. 230). Due to the passive nature of the translated instructions, Finnish participants reported their subjective experience rather than listing the specific suggestions they remembered (e.g., that their head was falling forward, that their eyes were closing). Researchers consequently scored such participants as highly amnesic. Similar results arose when translating and testing the HGSHS:A in a Swedish population; researchers calculated a 65% passing rate (Bergman, Tretter, & Kallio, 2003). They blamed the implausible outcome to translation problems as well, positing that participants misinterpreted the suggestion as “referring to changes in the content of consciousness as a result of suggestions delivered during the induction procedure. In the English original, the verb happen is used in the response booklet. However, it might be better to use the verb do in the translated versions” (Bergman et al., 2003, p. 354).

The next section explores why researchers cannot perform valid comparisons and reach legitimate conclusions without conceptual equivalence (Flaherty et al., 1988; Hunt & Bhopal, 2003).

**Conceptual Differences**

Dissimilarities in thinking that stem from unique sociocultural factors cause conceptual differences in languages, wherein similar words do not reflect the same construct in each language. In the case of post-traumatic stress disorder (PTSD), for example, both leading diagnostic manuals used in North American psychiatry—the Diagnostic and Statistical Manual, Fifth Edition (DSM-5: APA, 2013) and the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10: WHO, 2015a)—both emphasize perceived threats to life and bodily integrity. Some medical
anthropologists argue that what constitutes a threat, as well as how threats change from one society to another, and within societies, depend upon the experiences of people. For example, many of the classical instruments for measuring stress inputs in terms of life changes, such as the Social Readjustment Rating Scale (SRRS: Holmes & Rahe, 1967), overlook conceptual differences within societies. Young states:

> What was represented in the instruments that had been developed by the stress researchers was a translation of culturally normative ideas. . . . They weren’t applicable within . . . American society. . . . Much of the research was done in New York City and one of the highest stressful life events was being convicted of a crime and being sentenced to some kind of incarceration, it could be brief, but incarceration and conviction. For middle class Americans, this certainly was true. . . . On the other hand, research included poor Black Americans and poor Hispanic, mainly at that time, Puerto Rican Americans for whom incarceration was a less extraordinary event than it was for middle class White people. So there was . . . a gross difference between them.(A. Young, personal communication, September 26, 2013)

Similar cultural examples apply to the field of hypnotizability. For example, in Japan, researchers found that a direct translation of hypnotizability scales from English to Japanese resulted in significantly fewer words, thus considerably shortening the length of the induction procedure (Y. Fukui, personal communication by E. Sheiner, September 9, 2012). Japanese researchers also propose that a hypnotic experience is too holistic to be optimally measured using scales with such specific dimensions; they agreed upon the necessary establishment of new, more culturally appropriate scales.

Another example pertains to researchers who encountered validity problems when using a translation of the Diagnostic Interview Schedule (DIS) in Peru (Gaviria et al., 1984). Because North American researchers developed the DIS, its substance abuse section was not congruent with Peruvian culture. Many of the substances listed in the DIS were unavailable in Peru, while coca paste, a major drug in the country, was missing from the scale. Consequently, researchers drew on faulty DIS scores, resulting in inaccurate rates of Peruvian substance abuse. They encountered similar issues when translating the DIS into Hopi, a Native American language (Manson, Shore, & Bwosi, 1985). For example, one DIS item combines the concepts of guilt, shame, and sinfulness. Western populations often experience these feelings collectively; however, the Hopi strongly discriminate among all three concepts and typically experience them separately. Hopi participants therefore voiced their concern and asserted that the DIS necessitated three separate questions to correctly assess such a fusion of feelings in their population. The fact that certain items might not relate to the same normative concept in two cultures constituted another key problem. Auditory hallucinations, for example, are culturally consonant among Native Americans; the DIS, however, largely neglects this alternative cultural norm and incorrectly assesses it as a pathological symptom. Carlo Steirlin describes a similar diagnostic error in the case of a resident evaluating an aboriginal hunter from Northern Quebec. The patient presented with symptoms such as withdrawal, poor sleep, and loss of appetite, due to a preoccupation with the spirit of an animal he killed. He reported feeling “an estrangement of the spiritual connection between
himself and the fallen animal at the moment of the kill... which deeply disturbed him” (Steirlin, as cited in Guzder & Rousseau, 2013, p. 355). The resident diagnosed the patient with delusional disorder and prescribed a neuroleptic medication for treatment, thereby ignoring cultural explanatory models. In conclusion, in order to realize a meaningful translation of scales that acknowledges cultural significance, researchers must consider the diversity of unique meanings embedded in languages, such as:

The referents of symbols—i.e., their meaning—are aspects of a culture or a life world, not objects outside of language through which language obtains meaning. ‘Heart discomfort’ for Iranians is not the equivalent of ‘heart palpitations’ for Americans; it does not mean the same thing (Good, 1977). It is a symbol which condenses a distinctive set of meanings, a culture specific semantic network... (Good & Del Vecchio Good, 1986)

Without taking into consideration the effects of these sets of meanings, scales that follow Western biomedical standards cause misdiagnoses of cultural differences. Findings intimate that populations respond more or less strongly to hypnosis when, in actuality, they may understand the questions in the scales, or the entire concept of hypnosis, differently. As a construct specific to postmodern Western culture, hypnosis “cannot just be transplanted into other cultures” (Cardeña & Krippner, 2010, p. 743).

To decrease erroneous findings, hypnosis researchers would do well to allocate resources toward a comprehensive translation of local idioms and subtle linguistic complexities, and subsequently modify the original scales. Although this template may seem like a straightforward solution, researchers rarely follow it (Kleinman, 1987; McHugh & Slavney, 1986) and instead, often opt to use an asymmetrical translation procedure.

**Asymmetrical Translation**

Campbell and Werner (1970) define two categories of translations: symmetrical and asymmetrical. Symmetrical translation emphasizes the meaning, familiarity, and colloquialness of each language. Asymmetrical translation remains loyal only to the original language, which results in a translated version that “seems exotic and unnatural in the new language” (Jones & Kay, 1992, p. 187). Asymmetrical translation reduces the likelihood of conceptual equivalence between cultures and compromises the cross-cultural comparison of items (Jones & Kay, 1992). To attain scientific reliability, hypnosis scales must follow a symmetrical translation; test items must be common to all cultures under study for a proper translation of meanings and symbols (Campbell & Werner, 1970). A superficial translation of the original scales cannot attain symmetry because American psychologists designed and tested the items in the United States and developed their psychometric properties according to the English language. A symmetrical translation requires additional research regarding the cultural relevance of each test item and the creation of new test items. Nonetheless, researchers usually avoid tampering with and adjusting the test items according to the cultures under study (Chapman & Carter, 1979). The main reason for this reluctance remains the fact that achieving a perfectly
symmetrical translation is a complex and lengthy task. An asymmetrical translation, in
which researchers create a literal translation from the English hypnosis scales to the
language of the target population, requires less time and fewer resources.

Contemporary Research

Most experiments in which the researchers aim to disentangle cultural differences in
hypnotizability translate the American scales and apply them to a new population,
thereby ignoring issues of cross-cultural validity and reliability (e.g., Bergman et al.,
2003; Carvalho, Kirsch, Mazzoni, & Leal, 2007; David, Montgomery, & Holdevici,
2003; De Pascalis, Belluscio, & Russo, 2000; Lamas, Del Valle-Inclan, Blanco, &
Diaz, 1989; Lichtenberg, 2008; Lichtenberg, Shapira, Kalish, & Abramowitz, 2009;
Pyun & Kim, 2008; Roark, Barabasz, Barabasz, & Lin-Roark, 2012; Sanchez-Armass
& Barabasz, 2005). Some researchers do slightly modify the phrasing of test items;
when establishing Danish norms for the HGSHS:A, Zacharie, Sommerlund, and Molay
(1996) expressed that they changed some words to more appropriately attend to a Danish
audience. Due to the influence of a nation-wide famous stage hypnotist, who often
uses the words hypnosis and sleep in his performances, the researchers opted to use
trance or trance-state instead of hypnosis and sleep-like state instead of sleep (Zacharie,
Sommerlund, & Molay, 1996). However, most hypnosis researchers shy away from such
idiosyncratic changes and faithfully follow simpler translation procedures.

The most common procedure, as previously discussed, involves one bilingual scholar
translating the original scale, sometimes followed by as few as one other bilingual
scholar reviewing the translation. Finally, a certified interpreter performs a back trans-
lation. The small number of people involved in the process, in addition to their lack
of professional experience with hypnosis—the scholars are often volunteers recruited
on site—weakens the validity of such translations. Researchers spuriously consider
the translation successful when comparisons between populations yield similar results,
and the minor differences left over fall under cultural factors. Many studies also use
already translated versions of the scales that were created in the 1970s and 1980s,
when translation procedures were more lax. Researchers usually assess reliability using
the Kuder-Richardson (KR20) formula (Rubini, 1975). If the Cronbach’s Alpha reli-
ability coefficient is too low, researchers may recalculate the statistical analyses with
the discontinuation criterion (Weitzenhoffer & Hilgard, 1962). These practices reflect
faulty methodologies and paradigms, and vastly decreased accuracy of hypnotizability
prevalence rates across cultures and languages (De Jong & Van Ommeren, 2002).

Critiques

Few authors have thoroughly addressed the negative implications of hasty, and ulti-
mately defective, translations of scales and tests in research (Birbili, 2000; Temple, 1997;
Temple & Young, 2004). Arthur Kleinman (1987), a leading figure in the criticism of translation, states that the very essence of ethnographic research centers on translation. In failing to address the issues attached to translation, researchers risk perpetrating a category fallacy, or the reification of a nosological category developed for a particular cultural group that is then applied to members of another culture for whom it lacks coherence (Kleinman, 1977). Comparing the strictly Western concept of hypnotizability across cultures constitutes such a fallacy because hypnosis is “embedded in a network of implicit meanings and ontological assumptions about the nature of reality and of hypnosis itself” (Cardeña & Krippner, 2010, p. 746). For example, individuals in South American societies may experience a syndrome entitled soul loss, characterized by a feeling of fragmentation or dissociation of the soul after experiencing a trauma. The exact phenomenology of this depressive disorder does not appear in Western medical categories. In fact, psychiatrists consider soul loss a culture-related specific syndrome, defined as a syndrome that is “closely and significantly related to certain cultural features in [its] formation or manifestation of psychopathology” (Tseng, 2006, p. 565). A researcher could decide to operationalize the symptoms of soul loss, organize them into a questionnaire, establish reliability for use in Western society, translate the items to English, and apply the final product to an American population. Despite a rigorous translation, the data would lack validity because soul loss has no coherence for North Americans (Shweder, 1985). Such an assumption—that high reliability leads to high validity—constitutes a central mistake in psychiatric research (Kleinman, 1987).

**Validity**

Validity—the extent to which an instrument accurately measures what it purports to measure—requires a highly sophisticated approach and stands as one of the most difficult issues to face in transcultural psychiatry (Kleinman, 1987). In the specific case of measuring instruments such as scales, validity refers to the degree to which the tool measures what it claims to measure (AERA, PA, & NCME, 1999). High reliability, which indicates the consistency of observation over time, has proven more easily achievable in cross-cultural studies than high validity. A scale obtains a high reliability score if it produces similar results under consistent conditions (Carlson, Buskist, Heth, & Schmaltz, 2009). Validity requires understanding the particular cultural context (McHugh & Slavney, 1986). To illustrate this point, Kleinman (1987) used the example of 10 psychiatrists who were trained in the same assessment technique and who examined 100 Native Americans shortly after the participants had experienced the death of a family member. The psychiatrists determined with almost perfect consistency that the individuals reported hearing the spirit of the deceased calling to them, signifying a high reliability rating. However, the clinicians misattributed these self-reports of hearings to pathological auditory hallucinations, disregarding the fact that Native American tribes culturally encourage internal auditory experiences, rendering them an expected
phenomenon that hardly signifies mental instability or psychosis. To report these experiences as hallucinations, with the pathological weight that hallucinations connote in the medical field, is reliable but scarcely valid. Although Kleinman created this exemplary story to critique the DSM-III (APA, 1980), many scientists believe that the following versions of the DSM, including the most recent DSM-5 (2013), have yet to properly address the problem of cross-cultural validity (Phillips, 2013). Validity means more than the verification of concepts used to explain observations; it signifies the verification of the meaning of those observations in a particular socio-cultural system. Reliance on ethnography would help avoid major validity fallacies.

An Ethnographic Approach

Because the attainment of validity pertains to a fundamentally ethnographic enterprise, researchers studying hypnosis would benefit from collaborating with cultural anthropologists. To date, the fields of medicine and psychology have dominated research on hypnosis, with neuroscience and pharmacology meekly following far behind (see Figure 3). To many scientific researchers, observations directly represent reality (McHugh & Slavney, 1986). Through the eyes of a clinician, for instance, the word hallucination often points to an abnormal mental state; however, for the anthropologist,
it signifies a meaningful phenomenon in a world mediated by the cultural apparatuses of language, categories, and taxonomies. Western scientists seldom consider hallucinations as merely pathological in nature; some researchers, especially in the 1970s and 1980s, argued that individuals experience hallucinations on a continuum (Chapman & Chapman, 1980; Chapman, Chapman, & Raulin, 1976; Chapman, Edell, & Chapman, 1980; Claridge, 1985, 1987, 1990; Spitzer, Endicott, & Gibbon, 1979). In other words, these occurrences may range from innocuous and unexceptional (e.g., “The sounds I hear in my daydreams are usually clear and distinct”) to potentially pathological (e.g., “I have been troubled by hearing voices in my head”) (Bentall, 2000, p. 90). Thorough qualitative research and meticulous translation strategies can contribute to a stronger foundation and lead to a better understanding of hypnotic phenomena as they manifest across cultures (Jones & Kay, 1992).

A Model for Research in Transcultural Hypnosis Research

A standard ethnographic approach begins with the organization of focus groups, which involve group discussions among five to ten people from the relevant culture on topics proposed by a facilitator (Krueger, 1988; Morgan, 1988). Data collection relies on the interactions among the participants. Focus groups give insight into the sociocultural and political context of the population under study. By eliciting information on customs and identifying critical concepts, focus groups help researchers develop new test items and modify existing ones. Focus groups have several limitations (Morgan, 1988) that researchers can easily overcome (De Jong & Van Ommeren, 2002). For instance, focus groups tend to be unproductive when participants disagree on topics or feel uncomfortable talking about them. Pre-testing helps to determine whether certain topics will work in a specific focus group. To address concerns about the replicability of findings, researchers can run a sufficient number of focus groups per subpopulation, until the data generate consistent findings.

The next step in an ethnographic approach involves in-depth interviews, otherwise called person-centered ethnography (De Jong & Van Ommeren, 2002). These qualitative interviews elucidate subjective experiences and psychological processes affected by sociocultural factors. They provide background information that aids in the appropriate translation and adaptation of scales, and would ensure the relevance of the items.

The final step in preparing for the collection of transcultural data involves the organization of meetings in which researchers, colleagues, and representatives of the target population discuss the effects of culture and setting on the experience of participants. By integrating their conclusions into the research plan, researchers aim to link the design of the scales to the experience of participants and thus increase their relevance. Once researchers complete these steps, the hypnotizability scales will likely span a stronger anthropological foundation, bolstering both their reliability and validity.
Translation Procedures

The most important challenge in this ethnographic model comprises the translation and adaptation of the instruments while ensuring content, semantic, concept, criterion, and technical equivalences (Flaherty et al., 1988).

Content equivalence requires items to be relevant within the cultural context in use. For instance, in certain low-income areas of the world, parents may not be able to afford school fees; as such, asking whether a child attends school proves useless when diagnosing Conduct Disorder (De Jong & Van Ommeren, 2002).

Semantic equivalence signifies that the meaning of an item remains the same after translation. For example, with regard to PTSD, the word *nightmare* changes meaning when translated from English to Cambodian. In Cambodian, *nightmare* indicates a nightly visitation of a deceased family member, as opposed to a frightening dream in English.

Concept equivalence—the antonym of a category fallacy—requires that an instrument measures the same theoretical construct in both cultures. The aforementioned assessment of soul loss in the United States exemplifies a lack of concept equivalence.

Criterion equivalence means that the outcome of measurement of a variable matches another criterion, such as an independent assessment by a psychiatrist practicing in the culture under study. For example, healthy participants in African countries may indicate that they believe someone wants to harm them. In contrast to a Western psychiatrist, who might note symptoms of paranoia, a local psychiatrist may consider these statements normal if cultural customs include witchcraft and sorcery (De Jong, 1987).

Finally, technical equivalence entails a sensible administration of the instrument to avoid systemic biases. These biases may arise in several occasions, such as if the interviewer represents an emotionally loaded institution; if the setting lacks privacy; if the physical interpersonal distance proves to be culturally inappropriate; if social desirability encourages acquiescence rather than honesty; or if certain factors—such as gender, ethnicity, socio-economic status, or background—disturb the communication (De Jong, 1987).

Good ethnographic practice considers the contributions of both psychological and anthropological knowledge of qualitative data in such a way that they increase understanding of the cultural context (De Jong & Van Ommeren, 2002). Such approaches would enable easier validation of hypnotizability scales and improve interpretation of the results. The transcultural use of hypnotizability scales may benefit from following anthropological models, which boast more rigorous, systematic, and contextual approaches to the adaptation of instruments (Kleinman, 1987).

Conclusion

Every language carries deeply entrenched meanings inherent to its parent culture. An appropriate determination of a tenuous term such as hypnotizability would therefore
necessitate comprehensive prior knowledge of the cultural milieu in which the terms thrive beyond blind adherence to the psychometric parameters (e.g., reliability and validity) associated with the scales of choice. Behavioral scientists would stand to benefit from drawing on anthropological models to ensure that hypnotizability scales are culturally sensitive and socially appropriate.

Suitability of comparative measurements follows by grounding transcultural studies in local context through the organization of focus groups, in-depth interviews, and meetings with local leaders and members of the community. Moreover, by incorporating such considerations, we would be able to foster meaningful cross-cultural comparisons of hypnotizability, thereby permitting a more scientific understanding of the sociocultural factors that render certain populations more prone to hypnotic experiences than other groups, and the role such experiences play in specific communities. Scholars of hypnosis should work in concert with clinician-researchers (e.g., psychotherapists) and social scientists (e.g., medical and cultural anthropologists) to further unravel transcultural issues surrounding hypnotizability.

Acknowledgments

We would like to thank Mr. Mathieu Landry from McGill University (Canada) and Ms. Shelagh Freedman from Concordia University (Canada) for their constructive comments and edits on earlier versions of this manuscript, as well as Drs. Danielle Groleau and Allan Young from McGill University for their valuable insights and discussions.

Funding

Dr. Amir Raz acknowledges funding from the Canada Research Chair program, Discovery and Discovery Acceleration Supplement grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institutes of Health Research, and the Volkswagen Foundation (VolkswagenStiftung).

References


TRANSCULTURAL FACTORS IN HYPNOTIZABILITY SCALES


