

2011

A Cross-Cultural Adaptation of the Irrational Beliefs Inventory from English to Icelandic

Gudmundur Torfi Heimisson

University of South Florida, gheimisson@gmail.com

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#), and the [Psychology Commons](#)

Scholar Commons Citation

Heimisson, Gudmundur Torfi, "A Cross-Cultural Adaptation of the Irrational Beliefs Inventory from English to Icelandic" (2011).
Graduate Theses and Dissertations.
<http://scholarcommons.usf.edu/etd/3145>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

A Cross-Cultural Adaptation of the Irrational Beliefs Inventory from English to Icelandic

by

Gudmundur Torfi Heimisson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Psychological and Social Foundations
College of Education
University of South Florida

Major Professor: Robert F. Dedrick, Ph.D.
Kathy Bradley-Klug, Ph.D.
Yi-hsin Chen, Ph.D.
Carlos Zalaquett, Ph.D.

Date of Approval:
April 5, 2011

Keywords: Cross-cultural Psychology, Cross-cultural Measurement, Translations and Adaptations of Psychological Tests, Invariance Testing, Rational-Emotive Behavior Therapy.

Copyright © 2011, Gudmundur Torfi Heimisson

Dedication

To my parents

Acknowledgments

I wish to thank my major professor, Dr. Robert Dedrick, for all his guidance, encouragement, support, and patience. I would also like to thank my committee members Dr. Kathy Bradley-Klug, Dr. Yi-hsin Chen, and Dr. Carlos Zalaquett for their insights, suggestions, and feedback. Additionally, I wish to thank Dr. Herbert Exum for all his professional and moral support. Finally, I want to thank my former major professor Dr. Darrel Bostow and former committee member Dr. Melinda Hess for their contributions.

I also wish to extend my gratitude to the translation and adaptation team: Dr. Ellert Þór Jóhannsson and Dr. Kristján Guðmundsson who did the forward-translations, Dr. Anna Ingeborg Pétursdóttir and Dr. Sigurður Sigurðsson who did the back-translations, and Brynjólfur Borgar Jónsson, Ingibjörg Sigurjónsdóttir, and Margrét Dóra Ragnarsdóttir who were my cultural review committee.

Last but not least, thanks to the numerous faculty members in the U.S. and Iceland who let me into their classrooms to collect data, as well as to all the students who took part in the study.

Table of Contents

List of Tables	iv
List of Figures	vi
Abstract	vii
Chapter One Introduction	1
Purpose of the Study	7
Phase 1: Translation and Back-translation of the IBI	7
Phase 2: Cognitive Interviews and Cultural Validation of the IBI	8
Phase 3: Psychometric Analysis of the IBI	8
List of Definitions	8
Chapter Two Review of the Literature	10
Iceland	10
A Brief Introduction to Cognitive-Behavioral Therapy	13
Rational Emotive Behavior Therapy (REBT) and the Concept of Irrational Beliefs	15
Measurement of Irrational Beliefs	16
The Irrational Beliefs Inventory (IBI)	18
Cross-Cultural Research	21
Cross-Cultural Psychometric Research in Iceland	22
Translations and Adaptations of Psychometric Instruments	24
Initial Translation	24
Synthesis of Translations	25
Back-Translation	25
Expert Committee Review and Cognitive Interviewing	26
Pilot Testing	27
Submission of the New Document and Reports to the Developer of the Original Instrument	27
Psychometric Investigation of the IBI in the Present Study	28
Cognitive Interviewing and Other Qualitative Judgment Strategies to Support Construct Validity	32
Summary	35
Chapter Three Method	36
Phase 1: Translation and Back-Translation of the IBI	36
Forward-Translation and Synthesis of the First Icelandic Version of the IBI	37

Back-Translation of the IBI	39
Evaluation of Cultural Equivalence of the Icelandic Version of the IBI	40
Phase 2: Cognitive Interviewing and Cultural Validation of the IBI.....	41
Cognitive Interviewing	41
Phase 3: A Psychometric Analysis of the IBI.....	43
Procedure	46
Data Analysis.....	48
Chapter Four Results.....	49
Phase 1: Translation and Back-Translation of the IBI.....	49
Forward-Translation of the IBI.....	50
Cultural Relevance.....	511
Meaning	53
Back-Translation of the IBI	555
Cross-Cultural Equivalence of the IBI.....	588
Phase 2: Cognitive Interviews and Cultural Validation of the IBI.....	59
Cognitive Interviews and Discussion Group – U.S.	60
Cognitive Interviews - Iceland.....	64
Cognitive Interviews - Overview	66
Phase 3: A Psychometric Analysis of the IBI.....	66
U.S. Sample Statistics	68
Iceland Sample Statistics	722
Confirmatory Factor Analysis of the IBI	755
Model Fit.....	777
Factor Loadings	822
Residual Variances.....	85
Factor Correlations.....	877
Section Summary	89
Invariance Testing of the IBI	89
Chapter Five Discussion	98
Theoretical Background of the IBI and the Current Findings	98
Overview of Findings	101
Phase 1: Translation, Back-Translation, and Adaptation of the IBI.....	103
Phase 2: Cognitive Interviews and Cultural Validation of the IBI.....	105
Phase 3: Psychometric Analysis of the IBI.....	111
Limitations of the Study.....	113
Significance of Study and Recommendations for Future Research.....	115
References.....	127
Appendices.....	137
Appendix A The Irrational Beliefs Inventory - U.S. Version Arranged by Subscales.....	138

Appendix B The Irrational Beliefs Inventory - Icelandic Version Arranged by Subscales.....	140
Appendix C Recruitment notes to professors - U.S. and Icelandic Versions	142
Appendix D Informed Consent Note to Participants - U.S. and Icelandic Versions	143
Appendix E IBI - Survey/checklist for translators, back-translators, and cultural validators - U.S. and Icelandic Versions	144

List of Tables

Table 1. Qualifications of Translators and Adaptation Consultants	41
Table 2. Gender Distribution between U.S. and Iceland Samples.....	44
Table 3. Selected Examples of Items Flagged in Forward-Translation.....	51
Table 4. Items Flagged by Back-Translator 1.....	57
Table 5. Items Flagged by Back-Translator 2.....	60
Table 6. Items Flagged by Cultural Evaluators.....	58
Table 7. Items Identified as Problematic by U.S. Discussion Group	61
Table 8. Items Identified as Problematic in U.S. Cognitive Interviews	62
Table 9. Items Identified as Problematic in Icelandic Cognitive Interviews	65
Table 10. Percentage of Missing Data across Countries, Grouped by Item	68
Table 11. Item Descriptives and Item-to-Total Correlations for the U.S. Sample.. ..	69
Table 12. Item Descriptives and Item-to-Total Correlations for the Icelandic Sample.....	72
Table 13. Internal Consistency of IBI Subscales (Cronbach's α)	75
Table 14. Confirmatory Factor Analysis for the IBI, Icelandic and U.S. Samples	79
Table 15. CFA for the IBI, Icelandic and U.S. Samples, Modified Model	80
Table 16. Item Pairs with Correlated Errors on U.S. and Icelandic Versions of the IBI	81
Table 17. Unstandardized and Standardized Loadings for the Multigroup CFA	83

Table 18. Standardized Residuals Produced in the Multigroup CFA, by Country.....	86
Table 19. Current Correlations of Subscales Compared with Bridges & Sanderman (2002) and Koopmans et al. (1994)	88
Table 20. CFA Models Used in the Cross-Cultural Invariance Testing of the IBI	91
Table 21. Items Contributing to Difference Between the Icelandic and U.S. Samples When Factor Loadings Were Constrained to be Equal	92
Table 22. Items with Equal Loadings across the Samples, Arranged by Change in Chi-square	93
Table 23. Items Used in the Intercept Invariance Test, Arranged by Change in Chi-square	95
Table 24. Intercepts for the 21 Items Not Used in the Intercept Invariance Test Due to Unequal Factor Loadings	96
Table 25. Items on the IBI on which Equivalence was Tenable.....	113

List of Figures

Figure 1. CFA Model for the Icelandic Adaptation of the IBI	76
---	----

Abstract

The Irrational Beliefs Inventory (IBI) was built to measure self-defeating beliefs as conceptualized in Rational-Emotive Behavior Therapy. The IBI has five factors: worrying, rigidity, problem avoidance, need for approval, and emotional irresponsibility. A three-phase cross-cultural study was conducted to translate and adapt the IBI from English to Icelandic, and a Confirmatory Factor Analysis (CFA) approach was used for a test of factorial validity and cross-cultural invariance.

In Phase 1, the IBI was translated from English to Icelandic, using a forward-translation and back-translation. Two forward-translators and two back-translators were recruited. In Phase 2, qualitative interview methods were used in both the U.S. and Iceland to gain insights into the meaning of the items on the IBI. In the U.S., 21 university students provided insights in a group discussion, and four students were individually interviewed in depth about individual items on the IBI. In Iceland, four university students were interviewed in depth about the meaning of individual items. Three Icelandic psychology professionals were recruited to evaluate the appropriateness of the IBI for the Icelandic culture. In Phase 3, a Confirmatory Factor Analysis (CFA) was conducted to check factorial validity and cross-cultural invariance of the IBI. The total sample size in Phase 3 was $N=1547$, all college students, with $n=827$ in the U.S. and $n=720$ in Iceland. Overall, the CFA did not support the fit of the IBI's original five-factor model, although the fit was slightly better in the Icelandic version. Fit indices conflicted; the chi-square and comparative fit index (CFI) showed poor fit, while the RMSEA and

SRMR showed acceptable fit. Correlated error was found between 85 item pairs in the U.S. model, and between 68 item pairs in the Icelandic model. Modifications were attempted to the original model by including the correlated errors, and a multigroup CFA was conducted. Adding the correlated errors slightly improved the fit of both models, but only 11 out of the IBI's 50 items were found to have equivalent item factor loadings and intercepts between the countries. Results from the psychometric analysis and qualitative interviews indicated that the IBI needs to be rewritten if the measure is to be used for research in Iceland. The results were discussed in light of a recent analysis of REBT-based measurement instruments, and implications for cross-cultural research on highly abstract constructs such as irrationality were discussed.

Chapter One

Introduction

Cognitive Behavior Therapy (CBT) has been increasing in popularity worldwide since its emergence in the 1950s, and is now used in more than 30 countries worldwide. Rather than being a single form of therapy, CBT is an umbrella term for models of psychotherapy that commonly propose that a change in an individual's cognitions can lead to a behavior change (Dobson & Dozois, 2009). The idea that private cognitions could influence overt behavior contrasted with the earlier model of behavior modification, which asserted that only overt and directly measurable behavior was relevant to psychotherapy (Dobson & Dozois, 2009). Among the best-known approaches to CBT are Cognitive Therapy (CT) (Beck, 1967, 1976) and Rational Emotive Behavior Therapy (REBT) (Ellis, 1962). CBT approaches psychological suffering as a function of maladaptive learning and self-depreciation. Currently, CBT appears to be most popular in Western cultural regions such as the United States and Europe, although it is gaining popularity in Asia, Central America, and South America as well. It is difficult to ascertain exactly the extent of CBT's popularity. This could partly be because ideas and aspects of the approach are probably mostly tested in clinical settings and not necessarily documented in the literature, but also because CBT-related ideas potentially have a wider application outside of therapy and counseling settings, such as in education or cross-cultural research. The increasing popularity of CBT might indicate that it has struck a commonly human chord, but this cannot be assumed without careful testing of its methods and instruments in each specific culture where it is applied. For the purposes of

the current study, the main psychotherapeutic approach discussed in the current study will be REBT.

In general terms, Rational Emotive Behavior Therapy is an evidence-based, action-directed approach to psychotherapy that relies heavily on an empirically established knowledge base and a rigorous, continuing evaluation of progress from clients and therapists alike (Ellis, 1962). The approach was developed by Albert Ellis, who made the observation that humans have a unique capability to fear things that are not there, and may even never materialize (1962). This contrasts with other animals, most of which only show stress in the presence of a clear and present threat, and do not linger on the threat once it has passed (Ellis, 1962). Another fundamental assumption of REBT is that an interaction exists between people's perception of events and their psychological well-being; thoughts can affect self-perceptions and thus lead people to alter their framework of thinking (Ellis, 1962). Consequently, in times of distress and altered perception, the person suffers and learns self-defeating styles of thinking, marked by psychological distress. According to Albert Ellis, fears and anxieties can initially have roots in perfectly rational sources of stress that have been learned through sources such as classical conditioning and language (e.g., parental punishment, losing one's job, strains in relationships, etc.), but can then be mediated into other aspects of life due to the human animal's ability to associate symbols (e.g., words, images) with an emotional state. The central concept in the REBT framework is the construct of Irrational Beliefs. In Ellis's conceptualization, irrational beliefs are not psychological disturbances in and of themselves, but illogical and/or dogmatic beliefs that have turned into deeply held values through social conditioning and emotionally-oriented reasoning, which in turn go on to

cause emotional disturbances when a person's life experiences clash with those beliefs (Ellis, 1962). Ellis has described the process in terms of a system he called the ABCs of REBT, where A stands for either *Adversity* or *Activating event*, B stands for (*Irrational Belief*) about how things must or should be, and C stands for *Consequence* (Ellis, 1962). When one suffers adversity, real or anticipated, a normal and healthy consequence is to become upset, worried, or frustrated. However, when an adverse event is interpreted by someone with an *irrational* (absolutistic, rigid, etc.) belief about how things should or must be, the consequence is a maladaptive feeling such as anger, feeling of worthlessness, aggrandizement, or anxiety. Following from this reasoning, Ellis postulated that internal verbal reasoning mediates one's emotional view of oneself. As maladaptive ways of reasoning are identified during therapy, clients are taught how to dispute their self-defeating interpretations of events, dispute their irrational beliefs that caused those maladaptive interpretations to appear, and instead come up with alternate explanations, and thus de-condition themselves from their previously irrationally conditioned ideas (Ellis, 1962). From the general assumption that people can learn different styles of thinking based on their verbal explanations of events, it must follow that rationality is universally human. The spreading success of cognitive therapies indicates that the assumption has at least some merit. If the above are correct, then rationality and irrationality go hand in hand and should be universally exhibited across our species, regardless of cultural differences. However, if beliefs are verbal and symbolic, and a culture is a society defined by shared beliefs communicated verbally and symbolically, then it follows that psychological measurement instruments of verbal appraisal need to be appropriately adapted to the language of that particular culture.

From a research standpoint, ways to measure the constructs behind Cognitive Behavior Therapies have not always followed the rapid international progress of the clinical practice. Individual measures of psychopathology, for example the Beck Depression Inventory and Beck Depression Inventory II, have been translated and adapted into more than 30 languages, but measures of more abstract constructs such as rationality and irrationality have not been as widely developed and adapted. This may be partly because irrational beliefs are not a psychological disturbance as such, but a broad set of maladaptive verbal habits that *cause* a number of psychological disturbances (Ellis, 1962; Ellis & Harper, 1997). Another issue highlighting the broad nature of irrational beliefs is that attempts to measure irrational beliefs in the 70s and 80s seem to have yielded inconsistent results depending on whether the sample was clinical or student-based (Smith, Rausch, & Jenks-Ketterman, 2004). In a recent review of tests on irrational beliefs, Terjesen, Salhany, and Sciutto (2009) identified 14 measures of irrationality with English as the original language, and found challenges with many of them, most notably psychometric properties, but also issues relating to construct validity and conceptualization of irrational beliefs, among which was the question of whether a test measuring irrational beliefs should have items describing only irrational beliefs, or reactions to/consequences of those beliefs. The consensus in the literature, thus, seems to be that need for further development of irrational beliefs test is pressing, and the Irrational Beliefs Inventory (IBI) was chosen by the current author to explore ideas in further development of such instruments.

The Irrational Beliefs Inventory (IBI) was introduced in the Netherlands in 1994 (Koopmans, Sanderman, Timmerman, & Emmelkamp, 1994). The authors of the IBI

defined irrational beliefs as "unrealistic verbal reasoning processes by which external events are interpreted and through which emotional stress is mediated" (Koopmans et al., 1994, p. 15). The IBI is a 50-item self-report questionnaire consisting of statements to which respondents indicate agreement or disagreement on a 5-point Likert-type scale. Statements either reflect beliefs defined as irrational (e.g., Item 5: "I want everyone to like me"), or rational (e.g., Item 17: "Nothing is upsetting in itself – only in the way you interpret it"). The IBI consists of five subscales (Koopmans et al., 1994):

1. *Worrying* (12 items). This scale reflects items that measure the tendency to worry over potential future setbacks and misfortunes.
2. *Rigidity* (14 items). This scale represents the rigidity of one's values, subjective norms, and social norms. A high score indicates a strong tendency to feel or attribute guilt and blame for mistakes.
3. *Need for Approval* (7 items). This scale measures the need for others' approval, and fear of failure and rejection.
4. *Problem Avoidance* (10 items). This scale measures one's tendency to avoid problems, and to be dependent on others for decisions involving risk-taking.
5. *Emotional Irresponsibility* (7 items). This scale reflects the tendency to deflect attribution of emotions away from oneself and towards external causes.

The IBI was originally tested in Dutch, but ideas for the items comprising it came from two instruments developed in English in the United States; the instruments were the Rational Beliefs Inventory (RBI; Shorkey & Whiteman, 1977) and the Irrational Beliefs

Test (IBT; Jones, 1968). The RBI and IBT were found to be similar to each other in content, but performed poorly under psychometric investigations in the United States, their country of origin (Lohr & Bonge, 1982; Himle, Hnat, Thyer, & Papsdorf, 1985). When Koopmans et al. (1994) created the IBI, they combined the RBI and IBT to form a 137-item questionnaire, and translated the items to Dutch. After psychometric tests and modifications, the number of items was eventually narrowed down to 50, and the finished work was introduced in its final form as the Irrational Beliefs Inventory (Koopmans et al., 1994). In 2002, Bridges and Sanderman created the English version, which has been used in various studies, including investigations into the role of irrational beliefs in paranormal beliefs, and how irrational beliefs can affect procrastination (Bridges & Sanderman, 2002). For the U.S. version, the original Dutch version was translated, and reliability coefficients (α) of the five subscales were compared. Bridges and Sanderman (2002) found comparable results; internal consistency coefficients for the Dutch sample ranged from $\alpha = .71-.84$, while the internal consistency for the U.S. sample ranged from $\alpha = .69-.79$. Overall, internal consistency was slightly higher in the Dutch sample, except for the rigidity subscale, where the U.S. sample was slightly higher. For the current study, in light of calls for more international research with the IBI (Bridges & Sanderman, 2002; Koopmans et al., 1994), the U.S. version of the inventory was translated and adapted to the Icelandic culture, followed by an investigation of its cultural equivalence and comparison of its psychometric properties between the U.S. and Icelandic versions.

The cognitive-behavioral approach to psychotherapy – including REBT – is one of the fastest-growing counseling approaches in Iceland. Iceland has a professional society of CBT practitioners and several well-known measurement instruments related to

psychotherapy and personality assessment have been translated and adapted to the Icelandic culture. Examples of those instruments are the Beck Depression Inventory II (Arnarson, Olason, Smari, & Sigurdsson, 2008), the MMPI-II, and the NEO-PI-R. In spite of a well-established tradition for CBT in Iceland, more psychometric research is needed in the country to support both research and practice (Guðmundsson, 2005). By adapting the IBI to the Icelandic language and culture, the current study brings further understanding of the nature of irrational beliefs in general, and provides scientists and practitioners in Iceland with a measure to support the growing need for psychological services. From a broader perspective, the translated instrument can potentially help extend cross-cultural research on the construct of irrational beliefs.

Purpose of the Study

The overall purpose of this study was to fill the need for an instrument that could be used in Iceland to measure the key constructs in REBT, irrational beliefs. The creation of an Icelandic version of the IBI makes it possible to conduct cross-cultural research and evaluate the generalizability of irrational beliefs across samples in the U.S. and Iceland. The specific purposes of the study were to: (a) translate and adapt the IBI from the U.S. version to the Icelandic culture, (b) conduct a psychometric investigation of the Icelandic version of the IBI, and (c) evaluate the cross-cultural equivalence of the U.S. and Icelandic versions.

Phase 1: Translation and Back-translation of the IBI

In this first phase, the IBI was translated from U.S. English to Icelandic, first using a forward-translation, then a back-translation. Two forward-translators were employed, and the researcher also translated the IBI. Local Icelandic experience with

adaptations of psychological tests suggests that translators with a layperson's knowledge of psychology tend to use over-complicated language that confuses test takers (Jon Fridrik Sigurdsson, verbal communication, 2007). Two back-translators familiar with clinical practice were recruited for the study.

Phase 2: Cognitive Interviews and Cultural Validation of the IBI

Following the translation, a panel of three Icelandic evaluators, all of whom held degrees in psychology, was consulted to determine the cultural appropriateness and equivalence of irrational beliefs as measured by the IBI. The panel evaluated colloquialisms, sayings, and insights on variables that might not have come across in the translation, such as the tendency to express one's emotions. Other issues addressed included a qualitative evaluation about potential cultural differences between Icelanders and Americans about how they express their opinions about themselves, and the general relevance of the results to the Icelandic culture.

Phase 3: Psychometric Analysis of the IBI

In the third phase of the study, the psychometric properties of the U.S. and Icelandic versions, including internal consistency and factorial validity, were examined. For the psychometric analysis phase of the study, the factor structure of the IBI was examined using a confirmatory factor analysis, and comparisons of the factor structure across cultures were conducted using multigroup confirmatory factor analysis.

List of Definitions

- *Cross-cultural research.* An approach to discovering universals in human behavior by comparing different cultures. The idea is that, the more different

cultures show the same trait or behavior, the more behavior is likely to be common to all humans.

- *Cross-cultural translation of psychological tests.* A direct, straight translation of a test from one language to another. Considered a crude form of preparing a test for a new language and culture.
- *Cross-cultural adaptations of psychological tests:* An advanced form of test translation that takes customs, habits, vernacular, idioms, and other idiosyncracies of the target culture into consideration to make the test more sensitive to the new population for which it is intended. Involves employing committees of cultural reviewers, qualitative judgment techniques to determine meaning of items, and pilot-testing in the new language. A yet further advanced form of adaptation is called *assembly*, in which items are completely rewritten, or new items rewritten, to accommodate for subtleties in the target culture.

Chapter Two

Review of the Literature

The literature review is organized into eight sections. First is a brief overview of demographics and cultural variables in the target country, Iceland. The second section introduces Cognitive Behavior Therapy in general, and the third will focus on Rational Emotive Therapy (REBT) and the construct of irrational beliefs in particular. The fourth section will be about the development of the original IBI and currently available reliability and validity evidence on the instrument. Cross-cultural research will then be discussed in general, followed by a rationale of the procedures recommended to adapt measurement instruments for cross-cultural purposes, and how they apply to the current study. This will be followed with a discussion of cross-cultural adaptations and validations of psychometric instruments, using the adaptation of the IBI as an example. The last two sections will focus on two different approaches to validation which will be employed in the study, cognitive interviewing and psychometric analyses including multigroup confirmatory factor analysis.

Iceland

Iceland is an island in the North Atlantic, with a geographic area of 100,300 km² (62,687.5 mi²), and is the westernmost country in Europe (Statistics Iceland, 2011). The population in December 2010 was 317,630, about two-thirds of whom were living in urban settings (Statistics Iceland, 2011). The form of government is a parliamentary republic, elected every four years, with a Prime Minister as the head of government. The native and only official language is Icelandic, a North Germanic language of the same linguistic family as Faroese, Danish, Norwegian, and Swedish (Torp, 2004). Children

start first grade of primary school at six years of age and start studying English as a foreign language in fourth grade, but individual schools are allowed and encouraged to start teaching English between first and third grade (Ministry of Education, Science, and Culture, 2002). Danish is formally added in the seventh grade, but the ministry of education encourages schools to start precursory introduction to the language in fifth to sixth grade (Ministry of Education, Science, and Culture, 2002). In secondary school, a third language is added to the curriculum, the most common option being German or French, although schools are allowed to offer instruction in other languages if they have individual teachers qualified for teaching the required amount of third language credits.

Iceland has an economic and political model relatively similar to those of the other Nordic countries (e.g., Denmark, Finland, Norway, and Sweden), The economy is a social-democratic system with high taxes and an extensive welfare system that covers universal physical and inpatient psychological health care, although private medical and psychological practices exist for optional and/or non-essential services, and for referrals when demand is high on the state-covered system. Outpatient psychological services are almost entirely operated in the private sector, which is in contrast to the other Nordic countries, where such services are a part of the national health care system. Thus, confirmable statistics on the frequency of the use of outpatient psychological counseling are not easily accessible.

Icelandic society is relatively homogeneous, with around 94% of the population being of European origin, and most of the remaining 6% of various Asian descents (Statistics Iceland, 2011). Immigration to Iceland is consistently low and comes mostly from European countries, but the number of transients such as migrant workers and

foreign exchange students is much higher. In 2008, 247 people of other citizenships gained Icelandic citizenship, but 21,434 people (around 7% of the total population) from other countries were reported to be living in Iceland in 2008 (Statistics Iceland, 2011). In 2010, the total number of foreign citizens living in the country was 21,701 (Statistics Iceland, 2011). In 2009 and 2010, net emigration was 4,835, and 2,134, respectively. This contrasted with a more typical net immigration of 1,144 and 5,132 in 2007 and 2008, respectively (Statistics Iceland, 2011). The leading cause of the change in migration patterns was most likely the world financial crisis of 2008, which caused cataclysmic changes in Iceland's political and social systems. These included the temporary nationalization of all of the country's banks, International Monetary Fund intervention, and a resignation of the prime minister amid increasingly tense protests in the streets, followed by an interim election that produced a strong leftward shift in parliament from a political alignment that had leaned center-right relative to other Nordic countries since the 1980s (Boyes, 2009; Danielsson, 2009; UNCHR, 2010). Unemployment also rose and in 2009, a total of 8% of the workforce was registered unemployed, up from only 1.6% in 2008 and 1% in 2007 (Statistics Iceland, 2011). The financial crisis and its aftermath also affected quality of life as measured by international standards; in 2010, Iceland fell to 17th place in the UN Human Development Index, from being ranked third in the previous year (UNHDP, 2009, 2010). The complete effects of the crisis on the Icelandic population and culture have not yet been determined, and data on economic recovery are inconclusive at the time of this writing.

Iceland's closest cultural ties in areas such as foreign policy, arts and sciences, or student exchanges are with the other Nordic countries, although many students travel to

the U.S. for graduate school. General adult literacy levels have been estimated at 99% by the UN Human Development Programme (2010). In 2010, the population's net educational enrollment was around 97% in primary schools, around 90% in secondary schools, and 72% in tertiary education institutions (UNHDP, 2010). According to the 2009 report of the OECD's Programme for International Student Assessment (PISA), Iceland had 500 points on the average reading scale, which is statistically above the OECD average (OECD, 2010a), but a drop in seven points since 2000 (OECD, 2010c). According to the PISA study, less than 5% of the 4,410 students surveyed were socio-economically disadvantaged, and children of immigrant status did not perform in a significantly different way from native Icelandic children (2010b). In 2008, a total number of 1,637 book titles were published in Icelandic (5.1 titles for each 1000 inhabitants), out of which 1,217 were works originally written in Icelandic and 420 were translations from other languages (Statistics Iceland, 2011).

In Iceland, a professional society exists for psychologists, psychiatrists, psychiatric nurses, academics, and other professionals interested in cognitive and behavior therapies. The society is called Felag um Hugraena Atferlismedferd (FHAM), which could be translated to Society for Cognitive Behavior Therapy. According to the FHAM's web site, the society counts over 50 members, and regularly hosts continuing education events, lectures from international specialists, and seminars dedicated to cognitive-behavioral therapies.

A Brief Introduction to Cognitive-Behavioral Therapy

Broadly speaking, Cognitive-Behavioral Therapy (CBT) is an umbrella term for various schools of psychotherapy that propose that cognitions (e.g., thoughts, appraisals,

or beliefs) can affect changes in overt behavior, and thus by changing maladaptive cognitions into more adaptable ones, maladaptive behavior can in turn be changed (Abramson, Seligman, & Teasdale, 1978; Beck, 1976; Dobson & Dobson, 2009; Dobson & Dozois, 2009; Ellis, 1962). CBT started to take shape in the 1960s, largely due to widespread clinicians' frustrations with limitations of the dominant therapy forms of the time, most notably psychoanalysis and behavior modification (Beck, 1976; Ellis, 1962). Albert Ellis and Aaron Beck, progenitors of CBT and both former psychoanalysts, saw psychoanalytic theory as overly concerned with past experiences and hidden subconscious processes of unverifiable psychopathological impact (Beck, 1976; Beck & Freeman, 1990; Dobson & Dozois, 2009; Ellis, 1962; Ellis & Dryden, 2007). At the other extreme of the philosophical spectrum, Ellis and Beck found behavior modification techniques limited and simplistic due to the stringent methodology of working only with directly measurable overt behavior while deliberately ignoring less directly measurable covert behavior such as thoughts and beliefs (Beck, 1976; Beck & Freeman, 1990; Dobson & Dozois, 2009; Ellis, 1962; Ellis & Dryden, 2007). The solution was to adapt the position that private cognitions such as beliefs and private thoughts could indeed be accessed, and hence changed, by clients and their therapists when the client had been taught to identify and operationalize the relevant maladaptive cognitions and their behavioral consequences (e.g., Beck, 1976; Ellis, 1962). An active-directive clinical approach was gradually developed over time, for example by Ellis (1962), which entailed actively challenging a client's maladaptive cognitions during therapy sessions, teaching the client techniques to identify and argue against maladaptive cognitions as they happened, and giving the client homework assignments to reinforce those techniques

(Beck, 1976; Beck & Freeman, 1990; Dobson & Dozois, 2009; Dryden, David, & Ellis, 2009; Dryden, DiGiuseppe, & Neenan, 2003; Ellis, 1962; Ellis & Dryden, 2007).

Rational Emotive Behavior Therapy (REBT) and the Concept of Irrational Beliefs

Rational Emotive Behavior Therapy (REBT) is rooted in both ancient Stoicism and modern learning theory. On the one hand, REBT borrows the Stoic view that the best way to achieve happiness is through a realistic, self-directed, flexible, non-judgmental worldview (Dobson & Dozois, 2009; Ellis, 1962). On the other, REBT extrapolates from learning theory by proposing that words, ideas, and concepts can become classically conditioned stimuli that elicit an emotional response in the absence of the original unconditioned stimulus (Ellis, 1962; Skinner, 1971). In other words, humans not only possess the ability to respond emotionally to things that are not there, but also possess the ability of symbolic communication and language, which makes it possible to re-create an emotional event in memory, and even imagine an event that has never happened (Ellis, 1962). Responding emotionally to an event is not unhealthy if the event calls for it, as it is part of our species' ability to learn about the world, but the problem arises when an emotional response is exaggerated in the person's memory, and generalizes to future situations where the response is neither helpful nor needed (Ellis, 1962). Thus, if an emotion is repeatedly evoked without its appropriate situational context, it will interfere with natural ways of learning from experience, impede the person's happiness and survival, and hence become *irrational* (Ellis, 1962; Ellis & Dryden, 1987). From an REBT perspective, defining characteristics of an irrational thinking style includes a simplistic, rigid dogmatism and demandingness, along with what Ellis called "musturbatory" verbalizing (e.g., expectations verbalized as absolute demands),

unrealistic pessimism, low frustration tolerance, and overgeneralizing about oneself and others (Ellis & Dryden, 1997). From his clinical observations, Ellis initially compiled 11 summary statements that he believed represented latent causes of most psychological disturbances (Ellis, 1962). Ellis later revised and summarized the core irrational beliefs to three, which are stated along these general lines: 1) "I must be thoroughly competent, or I am worthless", 2) "Others must treat me considerately, or they are absolutely rotten", and 3) "The world should always give me happiness, or everything is pointless and worthless". In REBT sessions, clients are gradually taught to discern rational beliefs from irrational ones by observing how rigid demands and dogmatic interpretations of one's adversity lead to detrimental effects on their emotional well-being. Then, the clients are taught to challenge their irrational beliefs, for example by considering the evidence for and against the belief, considering the worst that could happen, the best that could happen, and substituting absolutistic phrases such as "I must", "I need", "it's awful", and "I'm horrible" for more adaptive phrases such as "I'd very much like", "I'd strongly prefer", "it's frustrating, but not unbearable", and "I'm fallible, yet unratable" (Dryden, DiGiuseppe, & Neenan, 2003).

Measurement of Irrational Beliefs. Despite the long history of REBT, and its apparently successful spread to different cultural areas, the body of literature on its effectiveness is scarce, and tends to consist of work done more by proponents of REBT than by objective researchers (Dobson & Dozois, 2009). Furthermore, research on instruments and tests measuring REBT constructs is also scarce, and only in the past decade has REBT theory been developing from a clinically focused theory into a personality theory with a social psychology component (Ellis, 2003a; Ziegler, 1999;

Ziegler, 2000). In a review of tests measuring irrational beliefs, published after the data collection for the current study had finished, Terjesen, Scituito, and Salhany (2009) identified and reviewed 14 tests measuring irrational beliefs with English as the original language, and found that the majority of existing REBT measures had conflicting or incomplete psychometric evidence, as well as lack of standardization (e.g., lack of formal manuals in many cases). A concern raised by Terjesen et al. (2009) was that many tests had items of questionable construct validity; while some items on some tests arguably measured beliefs, the wording of others represented constructs such as negative affect, maladaptive reactions, and psychological distress. Since the IBI was originally written in Dutch, it was not among the instruments reviewed by Terjesen et al. (2009), but had at the time been translated into at least two other languages from its original Dutch version (Bridges & Sanderman, 2002). The exact number of other non-English instruments of irrational beliefs is not known, but the emergent body of literature on irrational beliefs supports claims (e.g., Ellis, 2003a; Terjesen, 2009) that REBT is a theory in development, which in turn calls for further studies. It has been suggested that a possible explanation for REBT's underdeveloped status as a theory may rest within the REBT-CBT practitioner community, which may simply be more focused on clinical practice than on academic research (Ellis, 2004). However, any data-based and empirical approach to psychotherapy needs as much validation of its underlying constructs as possible, especially when the approach is applied in other cultures. If this is not done, there is a significant risk that the therapy may be culturally biased, and thus irrelevant and ineffective by default. Yet another explanation for REBT's lack of an international body of knowledge is that quantitative cross-cultural research is a relatively new field,

although it is decidedly on the rise (Harkness, Mohler, & Van de Vijver, 2003; Terjesen, 2009). Regardless of the reason for the lack of literature on the nature of irrational beliefs, universality of constructs should not be taken for granted when describing the human condition and applying solutions to perceived problems. Therefore, the need for more research on the instruments used to measure irrational beliefs seems apparent and urgent, and recent publications detailing advances in the operationalization of REBT constructs have indeed called for, and added to, the available literature to support such added research (e.g., McDermut & Haaga, 2009; Terjesen et al., 2009).

The Irrational Beliefs Inventory (IBI)

As Rational Emotive Behavior Therapy (REBT) (Dryden et al., 2009; Ellis, 1962; Ellis & Dryden, 1997) started to gain a foothold in the world of clinical counseling, several researchers attempted to build instruments to aid diagnosis and further assessment of irrational beliefs, but little consensus emerged (Ellis, 2004). Among the most widely used early research tools were the 37-item Rational Behavior Inventory (RBI) (Shorkey & Whiteman, 1977), and the 100-item Irrational Beliefs Test (IBT) (Jones, 1968). The IBI was developed by pooling and psychometrically analyzing items from the RBI and the IBT (Koopmans et al., 1994). The items from the RBI had initially been validated by administering them to professional clinicians at an English-language RET workshop supervised by Albert Ellis (Shorkey & Whiteman, 1977), but not much is known about the IBT items; the original publication of that test is a doctoral dissertation, and other data on those items do not seem to be available. The reason for pooling the RBI and IBT into a new instrument was that, even though those instruments – the RBI in particular - had high internal consistency, psychometric evaluations found them to have low construct

validity, measuring general negative affect rather than irrational beliefs per REBT theory (Bridges & Sanderman, 2002; Koopmans et al., 1994; Sanderman et al., 1987) .

The IBI itself was developed in the Netherlands, in Dutch, by Koopmans et al. (1994), over the course of three studies. In the first study, items were selected from the English-language versions of the RBI and IBT and translated to Dutch, but the translation and adaptation procedure have not been detailed. In the second study, the selected items were put to a further psychometric test to confirm the factor structure of the IBI. In Study 1, the RBI and IBT were administered to a two-group sample, a randomly selected Dutch non-clinical community sample that completed the surveys by mail, and a clinical sample of clients who had been diagnosed with social phobia. The non-clinical sample had 111 males and 120 females ($n = 231$), and the average age was 42.7 years ($SD = 13.1$, Range = 20-69 years), and the clinical sample had 33 males and 41 females ($n = 74$) of which the average age was 32.1 years ($SD = 8.9$, range = 18-56 years). An exploratory factor analysis was done to help decide which items to keep. For any item to be used for the IBI, a minimum loading of .40 was determined necessary. After the exploratory factor analysis, and qualitative judgment on items with loadings over .40, the number of items deemed usable for the original version of the IBI was reduced to 63 items from the original 137. The factor analysis yielded five factors that explained 36.3% of the total variance and were named Worrying, Rigidity, Problem Avoidance, Need for Approval, and Emotional Irresponsibility.

In Study 2, the new 63-item IBI was cross-validated with three other questionnaires, the General Health Questionnaire (GHQ) (Goldberg & Williams, 1988), the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975), and the Social

Cognition Inventory (SCI) (Van Kamp & Klip, 1981, see in Koopmans et al., 1994). This was done to help confirm the IBI's factor structure, and assess its discriminant validity by correlating its scales with scales on other instruments measuring traits such as neuroticism, anxiety, depression, and social dysfunction (Koopmans et al., 1994). The sample was a non-clinical community sample of 249 males and 200 females, with information about the gender of 36 people missing ($n = 485$). The average age was 42.3 years ($SD = 14.7$, range = 20-71 years). Multigroup CFA methods or IRT-based methods were not used to assess factorial invariance, but a Pascal-based computer program named PEKON (Camstra, 1985, see Koopmans et al., 1994) was used to measure the strength of potentially recurring factors, by comparing the correlation matrix of items in the new population to the rotated factor loading matrix of the items in the old population. The IBI-63 was then subjected to an exploratory factor analysis, with criteria for item elimination the same as in study 1. Based on this analysis, the number of items was reduced to 50. In study 3, the IBI-50 was put to the third exploratory factor analysis. The sample was a university student sample of 227 males and 307 females with gender of 4 participants unknown. Mean age of participants was 23 years ($SD = 5.2$, range unknown). The results of study 3 confirmed the results of study 2, and the IBI's 50-item version was confirmed as the final version.

The U.S. version of the IBI was tested for cross-cultural validity using a factor analysis with a varimax rotation, which resulted in the same factors as the Dutch version (Bridges & Sanderman, 2002). Other measures used to support the validity of the IBI were comparisons of Cronbach alpha coefficients of internal consistency, and correlations among subscales. It is not clear whether the original English-language items

that had been translated to Dutch to make the IBI were used in their original English form for the U.S. version of the IBI, or whether the Dutch versions of the items were translated to English. The U.S. and Dutch models were found to be identical (Bridges & Sanderman, 2002).

Cross-Cultural Research

In recent years and decades, cross-cultural research has been gaining in popularity in many academic disciplines, such as education and psychology, as a way of evaluating the generalizability of research findings in these areas (see e.g., Hambleton, 2005; Van de Vijver & Poortinga, 2005). Several journals now exist that are devoted in whole or part to the field of cross-cultural research, and the number of published studies in the cross-cultural field has been growing steadily over the past decades (Harkness, Mohler, & van de Vijver, 2003). An International Testing Commission exists, and its guidelines for international test use have been made available (ITC, 2001). In addition, several articles and chapters have been published in recent years to add to the existing knowledge base on needed steps for translations and adaptations (e.g., Hambleton, 2005; Sireci et al., 2006). Part of the reason for the recent added emphasis on cross-cultural research may be the fast pace with which information technology has evolved, and the extent to which it has changed the exchange of information in the world. What happens in one part of the world can now make a larger and faster impact in other parts than ever before, thus making the need for mutual understanding even more urgent. Ultimately, it could be argued that different cultural areas across the world have become no less dependent on each other than members of individual villages and towns were before the dawn of the information age. In addition to making the world smaller, technology has also provided

new opportunities to better understand human nature from examples of similarities and differences.

The field of cross-cultural measurement ranges from education to psychology to medicine. An example of a prominent field of cross-cultural research is academic achievement and other related variables. For example, the Programme for International Student Assessment (PISA) is an extensive international study on academic achievement and the role of cultural factors in issues related to schooling; variables investigated include scientific, mathematic, and reading literacy, along with attitudes, beliefs, and cultural variables believed to contribute to academic development (OECD, 2006; OECD, 2008). PISA started in 2000 and is conducted every three years, with an ever-growing number of participating countries; in 2006, 57 countries participated, and in 2009, 67 took part (OECD, 2008). Iceland has been a participating country in PISA since the first study was done in 2000.

Another example of a wide-ranging cross-cultural study in education is the Trends in International Mathematics and Science Study (TIMSS). TIMSS is an investigation of mathematics and science achievement in around 70 participating countries, and is conducted every four years (IEA, 2008). A sister study of the TIMSS, the Progress in International Reading and Literacy Study (PIRLS), is a similarly extensive study conducted tentatively every five years (IEA, 2008).

Cross-cultural Psychometric Research in Iceland. Fully adapted psychological tests are in short supply in Iceland (Guðmundsson, 2005). Several psychological tests have been translated into Icelandic over the last few decades, but most often, simple translation has been the option, without a formal standardization or adaptation, and many

psychological tests exist in Iceland that have not had their psychometric properties cross-culturally evaluated. Some of those tests have been used in the field with caveats, while others have been discontinued or scheduled for further testing. No complete list of translated and/or adapted instruments in Iceland exists, since the methods of translations have varied greatly, and the results have not necessarily been published (Friðrik H. Jónsson, personal communication, October 2009; Guðmundsson, 2005). Despite this, there are signs that the situation is improving. A formal call has been made in the Icelandic academic community to adopt the AERA (1999) standards and the International Test Commission's (2001) guidelines for translations and adaptations of tests in Iceland (Guðmundsson, 2005). In an Icelandic online library search, the researcher found several recent Icelandic peer-reviewed articles of tests that had been partially or fully adapted to the AERA and ITC standards and guidelines. In the case of published partial adaptations in Iceland, the authors usually qualify their work as pilot studies that require further and fuller adaptations, and used in the field with that caveat, due to an immediate and acute need for psychological instruments in Iceland. An example of such partially adapted tests is the U.S.-developed BULIT-R self-report measure of bulimic symptomatology (Thelen, Mintz, & Vander Wal, 1996), which was translated and back-translated, had measures of convergent and discriminant validity checks via other instruments, but went straight to local testing. Results revealed significant differences between clinical and non-clinical participants (Jónsdóttir, Þorsteinsdóttir, & Smári, 2005). Another example of a partially adapted test is the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kreammer, 1989). For this test, exploratory factor analysis, group mean scores, and reliability coefficients of subscales were determined and compared to more fully established versions in other

countries, with results indicating both similarities and differences between the Icelandic and U.S. versions. Further research was called for (Haraldsson, Smári, & Gylfason, 2001). Despite challenges posed to the Icelandic academic community due to the small size of the Icelandic-speaking population, fully adapted instruments in Iceland have been becoming more common. Examples include the first and second editions of the Beck Depression Inventory (Arnarson, Olason, Smari, & Sigurdsson, 2008), and the NEO-PI-R personality inventory along with its shorter version, the NEO-FFI-R (Costa & McCrae, 1992; Costa, McCrae, & Jonsson, 2002; Jónsson & Bergþórsson, 2004; Jonsson, 2005).

Translations and Adaptations of Psychometric Instruments

When adapting measurement instruments to new languages and cultures, it is necessary to ensure that the constructs at hand have an analogous meaning and function in both the original and new cultural contexts. When this has been done successfully, the result is called *construct equivalence* (e.g., Casillas & Robbins, 2005; Hambleton, 2005; Krishnakumar et al., 2004). The fluid nature of the terms "meaning" and "function" requires construct equivalence to be primarily established with qualitative approaches and judgmental strategies applied in a systematic way. Several authors (e.g., Hambleton, 2005; Sireci et al., 2006; Verra et al., 2006) suggest an approach to establishing construct equivalence that can be summarized into six basic steps which were followed in the current study. The steps are as follows:

Initial Translation. For the initial translation, it is generally recommended that multiple translators be used, and that all should be as close to bilingual and familiar with both cultures as possible, although it is cautioned that completely bilingual people may not exist (Sireci & Berberoğlu, 2000). Optimally, at least one translator should be

familiar with both the subject matter and testing procedures, and at least another should be used who is not familiar with either (Hambleton, 2005; Verra et al., 2006). To facilitate a process that achieves consensus without being too restrictive, structure in the form of clearly described goals is considered necessary (Wild, Grove, Martin, Eremenco, McElroy, Verjee-Lorenz, & Erikson, 2005). To achieve this, and to facilitate clear communication and have a common base from which to work with translators, the researcher constructed a form outlining the major goals of the translation, using both closed-ended and open-ended questions for translators (Appendix E). One caveat about bilinguals is that they may be a select sample, especially if the definition of bilingualism is complete fluency in both languages; however, this same potential drawback is that bilinguals can be utilized to respond to the same items on both the original and adapted versions of the test, both in relation to test-retest reliability and for examining differential item functioning (DIF; Sireci & Berberoğlu, 2000).

Synthesis of Translations. Once the translations have been obtained from the two independent translators, their versions of the measure should be compared by the researcher for any discrepancies. The translators should also be consulted on discrepancies.

Back-Translation. Once the instrument has been translated into the target language, it can be translated back into the original language, and the new back-translated version compared with the original. This is a procedure known as back-translation, and can be done to check for the conceptual accuracy of the original translation (Hambleton, 2005).

The back-translation approach is not without its detractors. For instance, Geisinger (1994) suggested that an editorial expert panel with qualifications comparable to the original translators might be more effective. Those experts might, for example, form their own opinion of each item and then compare notes on the survey item by item not much unlike when reviewing a journal article (Geisinger, 1994). In the current study, the difference between the back-translation approach and the expert panel approach was partly resolved by combining them; that is, arranging meetings with the translators together, as well as having meetings with an expert committee on the cultural equivalence of the items before and after the back-translation. Since the translators all lived in separate countries and states, this was accomplished by email. In addition, the researcher created a form for each translator and back-translator for troubleshooting and pinpointing of difficult items, which also facilitated discussion when meeting with a translator was not possible (See Appendix E).

Expert Committee Review and Cognitive Interviewing. Once the translation has been agreed on, a series of steps should be taken to ensure the relevance of items, from both the standpoints of test-takers and developers. From the developers' side, a committee of psychometric experts and/or people familiar with the target population's reading levels should judge the instrument's relevance to the target culture. This step relies on qualitative judgment strategies, and could be accomplished by having each committee member single out items that seem inappropriate for any reason, then have roundtable discussions on the corresponding items (Hambleton, 2005). A potentially important issue with translators related to bilinguality and dialect, is the locality of the translators' and back-translators' English. When studying English, most people in Europe learn British

English, which is not only different from U.S. English in spelling of words, but also has different meaning of many terms that can confuse adaptations (Koller, Aaronson, Blazeby, Bottomley, Dewolf, Fayers, Johnson, Ramage, Scott, & West, 2007). For the present study, this may not be a problem, since all the translators and back-translators had gone through higher education and/or professional development in the United States and were familiar with U.S. English and colloquial vernacular.

To evaluate the quality of the test items from the test-taker side, cognitive interviews with participants representative of the target population should be conducted (Beatty & Willis, 2005). A cognitive interview is the administration of a survey draft to individuals or small samples ($n \approx 5$) in order to gather additional verbal information about the survey items, usually by using specific probes and/or having respondents verbalize their thoughts (i.e., think aloud) as they are taking the survey (Beatty & Willis, 2005).

Pilot Testing. Before the test is administered, it is advisable to test the instrument on a small (e.g., $n = 30$) sample of test-takers. Factors gauged for could, for example, be education level, cultural values (e.g., if the test assumes knowledge of things not emphasized in the target culture), political spectrums (e.g., if concepts such as "left-wing" and "right-wing" correspond to comparable ideologies in the target culture), and familiarity with the test item format (e.g., Likert-scales, essay questions). Sources of bias related to the timing of the test also need to be looked into (Hambleton, 2005; Poortinga et al., 1994).

Submission of the New Document and Reports to the Developer of the Original Instrument. Once the new instrument is considered tested and ready to use, an accompanying manual with instructions on administration and scoring should be written

in the new language (Geisinger, 1994). It is recommended to notify the developer of the original instrument, as the new version will present new opportunities to add to the literature on its effectiveness and universality (Hambleton, 2005; Poortinga et al., 1994).

Psychometric Investigation of the IBI in the Present Study

A psychometric investigation of an instrument is a step towards validating the operational definition of the observed and latent variables that the instrument represents. Validity, in turn, is the extent to which theory and empirical evidence support the proposed interpretation of data (e.g., test scores) (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 1999; Crocker & Algina, 1986). Validity is defined by a set of procedures rather than a single definition; one proposed series of steps is to: 1) formulate a hypothesis about how the construct manifests itself, 2) select a measurement instrument representing solid operational definitions of the construct, 3) data collection, and 4) evaluation of the results (Crocker & Algina, 1986). From a psychometric validation standpoint, the primary focus of the current study is the factorial validity of the IBI's five-factor model (Koopmans et al., 1994; Bridges & Sanderman, 2002). Factorial validity represents the extent to which patterns (i.e., factors) observed in the data follow the patterns predicted by the relevant theory (Brown, 2006). In other words, factorial validity reflects how well a factor on a test represents the latent variable it is supposed to measure, and how unidimensionally (i.e., exclusively) the factor represents its target variable (Brown, 2006; Campbell & Fiske, 1959; Gerbing & Anderson, 1988). A common way to test factorial validity is Confirmatory Factor Analysis (CFA) (Brown, 2006). CFA can be used with one or multiple groups of

participants, and allows for a robust interpretation of the measurement model underlying the instrument, because it allows the researcher to create a statistical model to match the operationalization of the theory and examine all parameters of the model (Brown, 2006). Thus, given a good fit between theory and CFA coefficients, results of a CFA indicate high factorial validity and support for the current operationalization of the theory (Brown, 2006). In the case of a bad fit between the theory and the CFA, individual items, the whole test, or even the theory could be at fault (Brown, 2006).

When used with more than one group, a CFA can be used to help determine how well a theory generalizes across groups, and is then called a multigroup CFA (Brown, 2006). A multigroup CFA allows all parameters of a measurement model to be compared across groups. Multigroup CFA typically involves testing a series of models to determine sources of misfit. The degree to which the number of factors and their structure are invariant between groups is often called configural invariance (Brown, 2006). Invariance of the factor loadings themselves is called metric invariance. Invariance between groups on individual item intercepts is called scalar invariance, and invariance of error variances has been called strict factorial invariance (Brown, 2006; Meredith, 1993). Invariance between factor variances and factor covariances is called factor variance invariance and factor covariance invariance, respectively (e.g., Steenkamp & Baumgartner, 1998). Full consensus does not appear to have been achieved about invariance terminology at the time of the current study, but the terms above will be used interchangeably. A distinction has been noted between full and partial invariance (Brown, 2006; Meredith, 1993), which refers to the extent to which all levels of invariance are achieved. If not all levels of invariance are achieved, it is up to the researcher to examine the theory and the data to

determine what caused the test items or subscales in question to work differently (Brown, 2006; Steenkamp & Baumgartner, 1998). If there is noninvariance at the configural (i.e., factor) level, the results suggest that the latent variables on the test may not be represented equally across the groups. If noninvariance is demonstrated at the scalar (i.e., item) level, the results indicate what can be called Differential Item Functioning (DIF) (Teresi, Ramirez, Lai, & Silver, 2008). The presence of DIF on a test item can indicate bias in the test, but it can also identify that the test is sensitive to a quality or behavioral tendency that exists in one group and not the other, and is not accounted for in the original theory or test (Brown, 2006; Hambleton, 2006; Teresi et al., 2008). Different strategies exist to detect DIF, including the Mantel-Haenszel chi-square method (e.g., Holland & Thayer, 1988; Dorans & Holland, 1993), logistic regression (e.g., Hu & Triandis, 1985; Scott, Fayers, Aronson, Bottomley, de Graeff, Groenvold, Gundy, Koller, Petersen, & Sprangers, 2010), and methods based on item response theory (Lord, 1977, 1980; Lord & Novick, 1968). In a cross-cultural multigroup CFA approach, detection of DIF in an item tends to be included in the examination of metric (i.e., non-uniform DIF) and scalar invariance (i.e., uniform DIF). It has been suggested that IRT and CFA methods to detect DIF are compatible approaches (Stark, Chernyshenko, & Dragow, 2006), and while IRT and logistic regression methods have been successful in detecting DIF (see Teresi et al., 2008; Scott et al., 2010), it was decided to use multigroup CFA for the current purposes of the study. The use of the CFA approach is in line with Standard 9.2 for educational and psychological testing for situations when there is a reason to believe that language differences may cause two groups' performance to differ on a test (AERA, APA, & NCME, 1999). Multigroup CFA has been used extensively in cross-

cultural studies of measurement invariance, in fields ranging from education to psychology to medicine. An example of CFA in studies of education achievement include an in-depth analysis of variables involved in the TIMSS study (Wu, Li, & Zumbo, 2007). CFA has been used to test for invariance when tests have been translated from one language to another, including from Spanish to English (e.g., Furlan, Cassidy, & Perez, 2009), from Chinese (He & Wolfe, 2010; Wicherts & Dolan, in press). It has also been used to in multicultural samples within the same country. For example, Malcarne, Fernandez and Flores (2005) used CFA to examine invariance in a multicultural sample between Caucasian Americans, Filipino Americans, and Latino Americans on the Multidimensional Locus of Control Health scales. CFA is also used extensively in marketing and consumer research (e.g., Steenkamp & Baumgartner, 1998), and quite extensively in the medical literature (Acquadro, Conway, Hareendran, & Aaronson, 2008).

Despite lending itself well to cross-cultural research, psychometric testing using procedures such as multigroup confirmatory factor analysis is difficult, especially as the construct in questions becomes more and more abstract (e.g., from more narrowly defined constructs such as test anxiety to broader constructs such as personality; see e.g., Buss, 2001; Hopwood & Donnellan, 2010). REBT theory and its main construct of irrational beliefs have recently been moving more and more into the realm of personality theory. Given the abstractness and multidimensional nature of the construct of irrational beliefs, CFA is considered one of the best approaches to evaluate such a broad construct (e.g., Floyd & Widaman, 1995). Nevertheless, it is emerging that statistical procedures alone are not enough. Byrne and van de Vijver (2010) have specifically addressed this

issue in relation to cross-cultural testing, arguing that a blind statistical use of CFA is simplistic. Various researchers have called for holistic approaches to translations, adaptations, and testing of psychological instruments, and many recommend a mixed approach of statistical and judgmental methods (e.g., Brown, 2006; Hambleton, 2005, 2006; Rigdon, 1996; Sevigny, Savard, & Beaudoin, 2009; Solano-Flores, Backhoff, & Contreras-Niño, 2009). Thus, the current study followed the recommendations of Hambleton (2005), Sireci et al. (2006), and Beatty and Willis (2007) to use a mixed methods approach of quantitative multigroup CFA approach, and qualitative judgment strategies such as cognitive interviewing to aid and enhance the interpretation of the CFA fit statistics.

Cognitive Interviewing and Other Qualitative Judgment Strategies to Support Construct Validity

Although statistical procedures are valuable tools to help make sense of data, they do not by themselves reveal the sometimes subtle meanings underlying questionnaire items (Gigerenzer, 2004; Hambleton, 2005, 2006). Thus, any psychometric analysis of an instrument must go beyond the statistics and take the meaning of the relevant conceptual framework into account (Brown, 2006; Cronbach & Meehl, 1955; Rigdon, 1996). The most obvious approach is to consult the research literature on which the instrument is based, but other means have been applied, including cognitive interviewing, which is a qualitative approach to exploring an item's meaning for a respondent. Without an idea of an item's meaning in the sample, a statistical coefficient is of little use. Once the items on an instrument have been built, an important phase in development and validation is cognitive interviewing. Cognitive interviewing, as a judgment strategy, does not have a

formal definition nor is there consensus on the procedure that is universally agreed upon, but the broad definition entails administering draft questions to people or small focus groups in order to gather qualitative information about the survey items (Beatty & Willis, 2005). According to Beatty and Willis (2005), two main approaches to cognitive interviews have been used: "Think-aloud" and probing (Beatty & Willis, 2005). A think-aloud method entails having participants recite their thoughts out loud as they are completing a survey or test, thus allowing a researcher to observe how participants arrive at their responses, and whether or not items are being understood as intended.

Just as psychological measurement instruments need to go through adaptation procedures to be applicable across cultures, current evidence also suggests that the validation methods themselves might also need modification to be linguistically and culturally appropriate. For example, Pan (2004) conducted cognitive interviewing and found that when phrases commonly used for cognitive interviewing procedures (e.g., "how did you arrive at this answer?") were translated to Mandarin Chinese, the Mandarin-speaking respondents started to justify their answers rather than explain their thought processes. In other words, the wording and phrasing of questions, and possibly the research situation itself, evoked a task orientation with respondents more appropriate for a test than for a cognitive interview. In other words, as much as the constructs and wording on a measurement instrument may carry different meanings across cultures, so may the validation methods themselves (Pan, 2004). No published research exists on cognitive interviewing in the Icelandic language at the time of the present study, and comparative studies on cognitive interview methods have not been done formally in Iceland. However, email consultation with a U.S.-educated survey expert at Capacent

(formerly Gallup) Polling in Iceland suggests that the currently favored approach is to use simple translations of U.S.-based methods, and then to fine-tune the approach as the interview process goes on (Guðbjorg A. Jónsdóttir, email communication, June 2009). Levin et al. (2009) also suggest that focus groups or discussion groups can be useful along with one-on-one sessions with participants; if implemented appropriately, such a group could give several individuals a chance to discuss the meaning of an item. To avoid context effects and other kinds of contrived influence on participants' responses, a focus group format was not used in the present study for the group session. Instead, a more informal discussion group format was used. A potential drawback of such less structured approaches is that dominant members of the group might influence the others (Tourangeau, Rips, & Rasinski, 2000), but in the current study it became advantageous, as the purpose of the group discussion was to gain insight into the vernacular of the IBI as might happen in groups where dominant members exist naturally.

In the cognitive interviews themselves, guidelines from Tourangeau, Rips, and Rasinsky (2000) were adapted. Tourangeau et al. (2000) identified a model consisting of four stages of the process by which respondents in a cognitive interview might respond to a question. The benefit of such a model to the current study is that it provides a framework by which to conduct the cognitive interviews, as well as a helpful guide to evaluate the helpfulness of the information gained. The four parts of the model are comprehension, retrieval, judgment, and response. Tourangeau et al. (2000) then divide these four stages again into cognitive processes in which respondents might potentially be engaging in each stage. In the comprehension stage, the interviewee might attend to questions and instructions, try to make logical sense of it, identify the focus of the

question or the interviewer's intent, and try to link key terms to relevant concepts addressed. In the retrieval stage, the respondent will try to access the relevant information from memory, possibly use generic information of related or analogous experiences, and fill in the blanks from memory. In the judgment stage, the respondent might assess the completeness of his/her memories, draw inferences from available information, and make an estimate of his or her answer before giving it. In the final stage, the respondent will present the answer given, and possibly edit the answer if he/she has doubts about the one he/she gave (Tourangeau et al., 2000).

Summary

The purpose of the current study was to translate and adapt the Irrational Beliefs Inventory (IBI) from U.S. English to the Icelandic language and culture. The IBI's theoretical background is in Rational-Emotive Behavior Therapy (REBT), which conceptually belongs under the class of Cognitive Behavioral Therapies (CBT). The main focus of the IBI is irrational beliefs, which are a form of self-defeating thinking characterized by absolutistic demands, dogmatic thinking, and overworrying. The current study falls under the domain of cross-cultural research, and employs both quantitative statistical methods and qualitative judgment methods to evaluate how well psychological constructs are represented by different people of different social and/or ethnic groups, cultures, and/or countries. In the current study, Confirmatory Factor Analysis (CFA) and cognitive interviewing were employed and are discussed more fully in Chapters Three and Four.

Chapter Three

Method

The purposes of the study were achieved in three phases. In the first phase, the Irrational Beliefs Inventory (IBI) was translated from U.S. English to Icelandic. The second phase was a cross-cultural adaptation of the IBI. The third phase was a psychometric investigation of the Icelandic version of the IBI, followed by an invariance test to evaluate the cross-cultural equivalence of the U.S. and Icelandic versions.

Phase 1: Translation and Back-Translation of the IBI

The IBI's framework is grounded in Rational Emotive Behavior Therapy (REBT), which originated in the United States (Koopmans, Sanderman, Timmerman, & Emmelkamp, 1994). The IBI itself, however, was originally developed in the Netherlands and written in Dutch. To develop the IBI, new items were not constructed from scratch; instead, items were adopted and adapted from two instruments from the U.S., the Rational Beliefs Inventory (RBI; Shorkey & Whiteman, 1977) and the Irrational Beliefs Test (IBT; Jones, 1968, see Koopmans, Sanderman, Timmerman, & Emmelkamp, 1994), which were also rooted in the REBT framework. The final Dutch version of the IBI yielded a five-factor solution: Worrying (12 items), Rigidity (14 items), Need for Approval (7 items), Problem Avoidance (10 items), and Emotional Irresponsibility (7 items) (Koopmans et al., 1994). In 2002, a U.S. English language version was introduced, which produced the same five-factor solution with corresponding items on each scale (Bridges & Sanderman, 2002).

For the purposes of the current study, the U.S. version of the IBI was translated and adapted to the Icelandic culture. The reason for using the U.S. version instead of the

original Dutch version was practical; Icelandic experts on psychometrics who also speak fluent English are not hard to find, nor are good translators in general. However, it was judged to be difficult or even impossible to find an Icelandic psychometrics expert who also speaks fluent Dutch. The decision to use the English-language version was supported by the findings of Bridges and Sanderman (2002) who found the original Dutch version and the English-language version to be equivalent. The items on the English-language version are displayed in Appendix A, grouped by factors identified by Koopmans et al. (2004).

The back-translation method (e.g., Hambleton, 2005) was used to create the Icelandic version of the IBI from the English-language version, in the following steps to be explained in greater detail below: 1) Translation of the IBI from English into Icelandic by three translators, a process called forward-translation, 2) synthesis of the three translations into an Icelandic preliminary version with consultation from the forward-translators, and 3) translation of the Icelandic version of the IBI back into English by two translators who had not been involved with the forward- translation - a process called back-translation - to ensure that the meaning of the original version had been retained. After agreement had been reached between the researcher and back-translators on difficult items, the researcher constructed an Icelandic preliminary version, which was administered to a committee of cultural reviewers. These steps will be described in further detail below.

Forward-Translation and Synthesis of the First Icelandic Version of the IBI. Two forward- translators were recruited to translate the IBI into Icelandic (Table 1). The two forward-translators were Iceland-born and had Icelandic as their native language. The

first forward-translator held a doctorate in linguistics from a U.S. university. The second forward-translator had a doctorate in the history of psychology from a U.S. university, and also writes and edits psychology textbooks for Icelandic high-school students in the Icelandic language. A third forward-translation was done by the researcher himself. The researcher is a native Icelander whose native tongue is Icelandic, but who has lived in the United States for a total of about eight years when data collection commenced. To facilitate comparison and eventual synthesis of the three forward-translations, a comment form to help identify difficult items was emailed to the forward-translators along with the IBI (Appendix E). The researcher also used the comment form to make notes during his own forward-translation. Once all three forward-translations had been completed, the researcher compared them side to side on paper, and marked items that had been flagged as difficult to translate by any of the translators. A difficult item constituted any item marked as difficult to translate by any of the forward translators, or an item on which differences in meaning were detected between forward-translations. If an item had been flagged as difficult by a forward-translator on a comment form, the researcher consulted with the translator via email until an agreement had been reached on wording and meaning of all items. If wording of an item differed between translations but the meaning remained the same, the researcher chose the version that seemed to be clearest. Since the researcher and both forward-translators lived in three different countries when the study was conducted, email was the primary means of communication. The researcher assumed final judgment over any remaining differences between forward-translated versions, and synthesized the versions together into an initial Icelandic version of the IBI.

Back-Translation of the IBI. Once the initial Icelandic version had been synthesized from the three forward-translations, it was translated back into English, and the back-translated version was then compared with the original English version side by side, as a way of evaluating whether the items had retained their meaning in the translation process. The back-translation was accomplished with two Icelanders who both work as professors in the field of applied behavior analysis at universities in the United States (see Table 1). Since the researcher was familiar with the original English-language version of the IBI, he did not do a back-translation himself. Once the back-translations were completed, the researcher compared them with the original English version side by side on paper in the same way as the forward-translations. The purpose of this comparison was to check for discrepancies that might reveal shortcomings of the initial forward-translation, or to see if issues relating to shades of meaning needed to be addressed for the items in question. The back-translators were given feedback forms identical to the ones of the forward-translators, and were asked to flag any items they found difficult to translate for any reason. If comments on the feedback forms needed clarification, back-translators were consulted via email. Back-translated versions were not synthesized into one back-translated version, but were instead both compared individually to the original version. This was done because two back-translated versions were believed to be a more robust measure than one, and because the researcher's familiarity with the original version's wording could have diminished the effectiveness of a synthesized version. If a back-translator flagged an item as difficult, or when discrepancies were encountered between the back-translated versions and the original English version, each back-translator was contacted individually for feedback. The

researcher exercised final judgment over any changes made to the forward-translation as a result of the back-translation. No formal standards exist in Iceland about readability level, so the issue of readability level was left to the translators and cultural reviewers.

Evaluation of Cultural Equivalence of the Icelandic Version of the IBI. Even when an instrument has been translated so the words are the same, the instrument may nevertheless not be culturally equivalent across cultures, because the meaning of the words themselves may not be the same in the two cultures. Therefore, a qualitative analysis of cultural appropriateness was needed before a psychometric evaluation could take place. Among potential sources of cultural discrepancy in translation from English to Icelandic were the general tone of the items (e.g., whether the tone is consistent across the test, and whether the tone is appropriate for the topic), and whether all topics are equally relevant (e.g., role of religion in moral reasoning). For the cultural adaptation stage, three Icelandic consultants were recruited. All adaptation consultants had at least a B.A. degree in psychology, plus their respective graduate degrees in computer science, statistics and marketing research, and school psychology. The consultants and the researcher individually judged each item by its perceived cultural appropriateness, using a form constructed by the researcher to give reviewers structure and to facilitate communication later on, inspired by recommendations by other researchers (e.g., Hambleton, 2005; Wild, Grove, Martin, Eremenco, McElroy, Verjee-Lorenz, & Erikson, 2005). The form featured standardized questions on each item as well as on the wording in general. An open field was also included to allow comments not provided for in the questions from the researcher. In case of discrepancies between the consultants surrounding the Icelandic version of the IBI, the researcher conversed with the

consultants via email. The final version of the IBI was a synthesized version of the consultants' and researcher's text. Upon completion of this stage, Phase 1 was considered completed.

Table 1
Qualifications of Translators and Adaptation Consultants

Dimension of qualification	Forward Translator 1	Forward Translator 2	Back-translator 1	Back-translator 2	Adaptation consultant 1	Adaptation consultant 2	Adaptation consultant 3
Native language	Icelandic	Icelandic	Icelandic	Icelandic	Icelandic	Icelandic	Icelandic
Educational level	Ph.D.	Ph.D.	Ph.D.	Ph.D.	M.Sc.	M.Sc.	Cand. Psych.
Gender	Male	Male	Female	Male	Female	Male	Female
Field	Linguistics	Psychology/ History of psychology	Applied behavior analysis	Applied Behavior Analysis	Psychology (B.A.) Comp. Science (M.Sc.)	Statistics and Market Research	School psychology
Experience in instrument translation	None	Limited	Limited	Committee-work	Limited	Limited	Limited
Training in psychometrics	None	Licensed psych. counselor	Undergrad and grad courses	5 graduate courses	Standardized tests as graduate student	Extensive	Licensed school psychologist
Other experience	Icelandic Dictionary Vocabulary Staff	Psych. Counseling Psychology Textbook writing	Taught undergrad course in measurement	B.A. in English	Journalist Tech. Writer Tech. Transl.	Market research (10 years)	Licensed school psychologist

Phase 2: Cognitive Interviewing and Cultural Validation of the IBI

Before a full psychometric analysis could take place, the IBI needed to be pretested in both the U.S. and Iceland to spot problems that might occur during the full-scale administration. The pretesting was done with cognitive interviews in a small group and in one-on-one sessions in the U.S., but only one-on-one sessions were used in Iceland, due to time constraints of the study in Iceland.

Cognitive Interviewing. Since the U.S. version of the IBI had already been developed and was not to be changed for this study, cognitive interviews in the U.S. primarily served the purpose of identifying potentially problematic items. In the U.S.,

cognitive interviewing was done in two stages: an informal focus group stage, and an individual cognitive interviewing stage. The informal focus-group session was held with 21 students in a classroom. Specific demographic data were not collected for the group, but the participants were students in a class taught by the researcher. The class was on applied social sciences and was taught in the College of Education at an urban research university. The class is an exit requirement for non-education majors, taken by students from every college of the university and recommended for third and fourth year students. Students' participation was part of a classroom activity related to curriculum on psychometric instruments and their development. Participants were asked to respond to the items on the IBI and make notes of any items that were viewed as odd or unclear for any reason. When all group participants had finished responding to the IBI, the researcher read the items aloud to the group item by item and asked the group as a whole for a show of hands if they had a comment about each item. This was done to get a feel for potentially problematic items. Then, participants were asked for their comments on the items, and group discussions followed if participants disagreed.

The one-on-one cognitive interviews in the U.S. were done with four university students. All four students were female, aged 21-35 years old, and were former students of the researcher, recruited through personal communication. Two of the students were psychology majors, one was an environmental science major, and the fourth was a biology major. Similar cognitive interviews were conducted in Iceland with four university students in one-on-one sessions. The Icelandic participants were two males and two females. The males were a 24 year-old education major and a 36 year-old psychology major, and the females were a 24 year-old business major and a 25 year-old

law major. Recruitment of the four participants from Iceland was accomplished through personal communication (a group discussion session was not possible in Iceland due to time constraints on the researcher's stay in Iceland).

A combination of think-aloud techniques and follow-up probes was used. Participants were in advance told that there were no correct or incorrect answers to the survey, and were asked to answer the questionnaire as they usually would, but read the items aloud and think aloud as they decided which answer option to take. Participants were then instructed to ask the researcher if they did not understand an item. The researcher made notes on paper if participants hesitated, asked a question, or expressed an opinion about any particular item. After the whole IBI had been answered, the researcher asked probe questions about items on which participants commented (e.g., "did anything strike you as odd with this survey?", "how did you come up with this answer?" "What did you mean by "strange question in item X?"). The reason for conducting the pretesting procedures in this order was to explore participants' opinions as freely as possible, without possible interference from the researcher. The verbal probing focused on items that seemed to cause confusion in the participants as observed by the researcher, as well as those items in which the translators or consultants had disagreements during the forward-and back-translation stages. At the end of the interview, the researcher debriefed participants on the purposes of the IBI, restated an assurance of confidentiality, and noted any final thoughts participants might have had.

Phase 3: A Psychometric Analysis of the IBI

A total of 1582 university students in the U.S. and Iceland took part in the psychometric phase of the study, 849 of whom were from the U.S. and 733 from Iceland.

Age in the Icelandic sample ranged from 19-61 years of age with a median age of 24 years, a mean age of 26 years, and a standard deviation of 6.15. Qualified U.S. participants ranged from 18-61 years of age, with a median age of 21 years, a mean of 23 years, and a standard deviation of 6.59. Participants above 64 years of age were required to be excluded, but only one participant met that criterion, a 67 year-old in the U.S. sample. Both samples were approximately two-thirds female (Table 2).

Table 2
Gender Distribution Between U.S. and Iceland Samples

Gender	U.S.		Iceland	
	N	%	N	%
Males	208	24.5	223	30.4
Females	634	74.7	506	69.0
Missing	7	0.8	4	0.5
Total	849	100	733	100

In the U.S., recruitment was done exclusively in undergraduate and graduate courses taught in a College of Education in a state university in an urban area. The university is located in Florida in a city with around 300,000 inhabitants and serves a greater area of around 1.5 million. U.S. recruitment was done both by contacting professors face-to-face and by sending an emailed request for recruitment (See Appendix C). In the U.S., 21 professors were contacted, all by face-to-face requests and an email follow-up. The researcher started recruitment by contacting professors who taught large classes, then moving on to those who taught smaller classes. Of the 21 professors contacted, 19 provided permission to use their classes in the study, with two professors unable to participate due to time constraints of their course schedule.

In Iceland, recruitment was also done in an urban research university setting, but in two universities, one public and one private. The colleges in which data were collected were colleges of education, psychology, social science, public health, and business. The

Icelandic universities are in a city with a population of around 130,000 people and serve a greater metropolitan area of around 200,000. For recruitment purposes in Iceland, the researcher sent recruitment requests via email ahead of his trip to Iceland. Initially, professors were selected on the basis of class sizes, and professors teaching large courses were contacted first. The recruitment request was in the form of a personalized email, which was followed up in person once the researcher arrived in Iceland. At the public university in Iceland, class attendance was much lower than expected, which prompted the researcher to seek a backup sample. This was accomplished in three ways: firstly, professors with whom rapport had already been established were asked to recommend a colleague to contact; secondly, office staff at the respective department offices at the Icelandic university were asked for contact information of professors at the department; thirdly, professors at a private university were recruited. At the private university, the dean of a college of public health and education was contacted with an email request, and the dean personally helped the researcher by encouraging professors at the college to offer their classes. The demographics of the two Icelandic universities were deemed similar enough to warrant the use of the sample from the private university. A total of twelve professors in Iceland were contacted, all of whom gave permission for their courses to be used in the study. It should be noted that Icelandic universities have either 3- or 4-year undergraduate degrees depending on different colleges and schools, but do not classify their students as freshmen, sophomores, juniors, and seniors as is done in the U.S.

The three universities are different in size and ethnic diversity; the university in Florida has about 34,000 undergraduate students and around 8500 graduate students,

whereas the Icelandic public university has only about 10,000 undergraduate students, with around 3100 graduate students and 400 doctoral students at the time of this study. The private university had a total enrollment of 2872 students enrolled for the 2009-2010 school year, of which master's and doctoral-level students were 608. The Icelandic universities both have a number of foreign-born students, but most students are Iceland-born. Due to potential confounds resulting from taking questionnaires in a second language, only responses from U.S.-born students with English as a native language were deemed eligible in the U.S. sample, and only Icelandic-born students with Icelandic as a native language were eligible for the Icelandic sample. To simplify the administration procedure, ineligible participants were excluded after the questionnaire was administered. This was accomplished by including an extra item on the IBI demographics section in each country, requiring participants to state whether or not they were native speakers of Icelandic or English, respectively. Exclusions were based on citizenship and age. In Iceland, seven participants were of non-Icelandic origin and were thus excluded. Additionally, six participants did not state their citizenship, and they were excluded as well, leaving the Icelandic total sample at 720. In the U.S., one participant was 67 years old and thus excluded. In the U.S. sample, non-citizens were 14. Seven participants were excluded because they did not state their citizenship, leaving the total eligible U.S. sample at 828.

Procedure

The administration procedure was analogous between the U.S. and Iceland. The researcher administered the IBI personally in university classrooms in both Iceland and the U.S. The researcher usually arrived at the beginning of the class, in order to minimize

end-of-class distractions and maximize participation. Upon arriving in the classroom, the researcher would ask for a student volunteer to start distributing the IBI around the classroom. Occasionally in both Iceland and the U.S., the professor would help distributing and collecting the IBI among students. For a particularly large class of around 300 students in Iceland, the researcher brought two assistants he trained to facilitate distribution and collection, but this was not done in the U.S. While the IBI was being distributed, the researcher gave a short introduction of the study. When all participants had gotten a copy of the IBI to which an informed consent slip had been stapled, the researcher would read out the informed consent statement and ask participants to tear off the informed consent slip if they consented to take part. Once the informed consent statement had been read, the researcher read aloud the directions on the IBI itself, after which participants could begin answering the survey. The instruction phase took between 1-2 minutes. Participants were instructed to raise their hands if they had any question about wording or content, and to raise their hand when they were finished. If a participant did not understand an item once administration was in session, the researcher instructed the participant to answer the item in the way that made the most sense to him/her, and reminded him/her that there were no correct or incorrect answers. When a participant had finished, the researcher would come and pick up the IBI. Once all participants had finished, the researcher briefly stated the purpose of the study and thanked the professor and students for their time. Time taken to answer the survey in both countries ranged between 15-20 minutes, but time taken in each classroom only exceeded the 20-minute mark once, in a psychology classroom where the professor asked the researcher to stay and explain the study in greater detail. All professors received a thank-

you email from the researcher after the study, along with an attached letter to participants detailing the purpose of the study as well as the researcher's contact information if needed. Several participating professors in Iceland posted the letter on their course web sites. It was not possible to monitor exactly how many students refused to take part in the study, since the surveys passed out to each class were not counted. However, no overt protests were observed, and no abnormally fast responses were observed.

Data Analysis

Descriptive statistics were obtained for both versions of the IBI using SPSS v.18 statistical software (SPSS Inc., 2010). Most items on the IBI are worded in an irrational direction, but several items were worded in a rational direction, so they needed to be reverse-coded in SPSS. Confirmatory factor analyses for each country, intraclass correlation coefficients, and invariance tests between both countries were obtained in Mplus, versions 3 and 4.

Chapter Four

Results

The purposes of the study were to translate the Irrational Beliefs Inventory (IBI) from U.S. English to Icelandic, adapt it to the Icelandic culture, and then assess the cross-cultural equivalence of the U.S. and Icelandic versions. This was accomplished in three phases: Phase 1 was a translation and back-translation of the IBI from U.S. English into Icelandic, Phase 2 was a qualitative cultural validation of the new Icelandic version using cognitive interviews and other judgment strategies, and Phase 3 was a psychometric investigation of the IBI. The psychometric investigation included a confirmatory factor analysis of both U.S. and Icelandic versions, as well as an invariance test to assess their cross-cultural equivalence. This chapter is organized into three main sections, each detailing results obtained during the three respective phases.

Phase 1: Translation and Back-Translation of the IBI

A total of three forward-translations were produced, two back-translations, and three qualitative cultural equivalence reports from three cultural consultants. A comment form with four questions, created by the researcher, was provided to each translator and consultant. The purpose of the form was to provide structure for the reviewers, as well as consistency of responses to aid interpretation. In addition, an open comment field was provided on the form to allow feedback on issues of which the researcher might not have thought.

The form had the following items:

- Name: _____
- How long did your work on the IBI take you? Approx. _____ hours.
- Did you do your work in one session, or in several sessions? If several, please explain:
- Did any items raise flags because of strange/archaic/culturally inappropriate wording or that you would not expect to see in a psychological test? *Please state item # and concern below:*
- Did you encounter any items that were particularly difficult to translate into English, and if so, why? *Please state item # and reason below:*
- Any other thoughts/concerns/difficulties/comments:

A copy of the comment form is provided in Appendix E. The form was delivered to the review committee as an attachment via email, and delivered back to the researcher via email after the work was done. Informal email correspondence was also used to clarify comments and get additional judgmental data from translators and cultural consultants when needed. Qualitative data on problematic items produced by this method are presented in greater detail in the text and tables below. Each table has comments from all forward-translators (including the researcher), back-translators, and cultural consultants, respectively, but most comments were too wide-ranging to fit into the table format, and are discussed individually in the text below each table.

Forward-Translation of the IBI. Most forward-translators' comments were of a general nature (e.g., "this sounds weird", "I found the tone of this one difficult") and only a few specific comments were given on individual items in the forward-translation round. Instead, however, several comments were obtained regarding the flow and tenor of the

IBI, and the general tone of the instrument was found to be quite problematic to translate accurately into Icelandic (Table 3).

Table 3
Selected Examples of Items Flagged in Forward-Translation

Item # and English wording	Icelandic wording	Difficulty/disagreement
3. Certain people are bad or wicked and should be severely punished for their sins.	Sumt fólk er vont og illgjarnn að eðlisfari og það verðskuldar þunga refsingu fyrir misgjörðir sínar.	"Certain" is too specific, "some" might fit in better. Icelandic word for "misdeeds" is better here, but item will change meaning. Religious context not applicable to Iceland.
4. People should observe moral laws more strictly than they do.	Fólk ætti að virða almenn siðaboð og siðferðisleg gildi betur en það gerir.	"Moral laws" is stronger than "moral values". "Moral laws" sounds strange and unnatural in Icelandic.
46. People make their own hell within themselves.	Eymd og vanlíðan fólks er undir fólki sjálfu komið.	A reference to hell in Icelandic sounds very metaphysical, does not capture the tone of English version.

Sources of disagreement between the three Icelandic forward-translators (which included the researcher) were analyzed by comparing translator comments on sheets of paper side by side, and then writing down a qualitative judgment on the topic and disagreement at hand. Translators were contacted via email if clarification was needed, but otherwise the researchers' judgment and Hambleton's (2005) guidelines for translation and adaption were the primary means of content analysis and decision-making. The qualitative content analysis revealed two general categories into which disagreement or difficulties in translation could be analyzed: cultural relevance on one hand, and meaning on the other.

Cultural Relevance. Forward-translators offered both general comments concerning the overall tone of the IBI, as well as more specific comments about words or whole items on the IBI. Both forward-translators, both of whom attended graduate school in the U.S., remarked that they found the IBI's whole tone and topic range to have an

"American" feel to it. Specifically, one forward-translator remarked what translates to: "Many questions about moral issues, [sic] all a bit American in my opinion." The researcher and both forward-translators all agreed that, had the IBI been translated directly into Icelandic, it would have had a strong religious overtone, and would as such have sounded foreign in Iceland's secular culture. Therefore, to make a more culturally relevant version of the IBI, the researcher's goal during the synthesis process was to tone down the religious component and aim for a secular rather than religious tone where applicable.

Several specific terms, idioms, and items got singled out as potentially irrelevant, most prominently "approval" and "acceptance," which in U.S. English can refer to self-esteem concerns, but in Icelandic refer more commonly to an agreement about a specific topic, or approval of a specific act. Additionally, when translated directly, both terms translated to the same word, "samþykki" ("agreement"). To retain meaning of the relevant items, the term "approval" and "acceptance" were, depending on context, adapted to "umhyggja" (care), "viðurkenning" ("acknowledgement"), "virðing" ("respect"), and the phrase "falla öðrum í geð" ("be liked").

The term "moral laws" was found to be inappropriate in that context in Icelandic, since the word "law" either means "lög", which refers to written laws of the land, or "lögmál" which in Icelandic can mean "rule" or "natural law." To avoid associations with legal issues or natural laws, context-based terms approximating "general codes of conduct" and "dos and don'ts" were used.

The phrase "bad sexual conduct" raised disagreements among forward-translators. One translator translated "guilty of bad sexual conduct" to "sekir um vonða kynferðislega

hegðun", which suggests malevolent and possibly illegal intent, the other translated "bad" as "glannalega" which translates to "reckless", and the researcher initially translated it as "óábyrga" which translates to "irresponsible." In the end, the phrase "hegðar sér illa" was settled upon, which can refer to recklessness, carelessness, irresponsibility, or malevolence all at once depending on the context.

Meaning. Several English terms were flagged on the IBI that may have an equivalent in Icelandic but mean something altogether different if translated directly. One difficulty in adapting such terms comes with grammar; the Icelandic language often relies on action verbs and context, rather than nouns, to convey shades of meaning. In addition, the same word can often mean many things, depending on the context. As a result, a complete rewrite of an item may be needed to convey a change in tone, and this was indeed often the case when forward-translating the IBI. The following words and phrases generated the biggest flags or disagreement among the forward-translators.

Forward-translator 1, a psychological counselor, listed general difficulties with words such as "concern", "worries", "upset", "disturbed", and "unpleasantness," because he found them all to boil down to one word, "áhyggjur", the Icelandic word for "worries." In these cases, complete rewrites of the items were necessary to retain the original meaning of the item. Conversely, the English terms "like" and "dislike" have many equivalents in Icelandic, each having different context-based shades of meaning.

Several idioms on the IBI needed to be adapted to Icelandic equivalents. "Feeling blue", common in U.S. English to describe feelings of dejection and sadness, is an example of such an idiom. On the IBI, it appears in the item "A person won't stay angry or blue long, unless he/she keeps himself/herself that way." The color blue has no relation

to sadness in Icelandic, and a literal translation would suggest that anger and blue skin color last only as long as people choose. The closest Icelandic equivalent to "blue" in this context, was found to be "einmana", translatable to "lonely."

In the item, "People make their own hell within themselves", the word "hell" was replaced with the Icelandic words "eymd" and "vanlíðan", for "misery" and "malaise", but the meaning was believed to be preserved, since the item ostensibly measures the degree to which a person believes that emotional misery is self-inflicted. The Icelandic term "helvíti" ("hell") is a commonly used expression of frustration, analogous to "damn it" in English, and can in the grammatical definitive refer to a strongly disliked person (analogous to "the bastard" in English.)

Yet another example of a cultural restructuring during synthesis was the item phrased "Certain people are bad or wicked and should be severely punished for their sins." In Icelandic, the wording and tone of the item changed to "Some people are mean and malicious and they deserve harsh punishment for their wrongdoings." All translators agreed that the word "sin" might not relate as intended to Icelanders, because Icelandic translation, "synd", has a dual meaning and is often used as Americans would use "shame" or "pity". For example, an American might respond to spilling a glass of milk by saying something like "such a shame that I spilled all this milk," or "what a pity that I spilled all this milk," whereas an Icelander's take on the incident might translate to "such a sin and a shame that I spilled this milk". Yet, despite the word "sin", the phrase has no religious meaning at all when translated to Icelandic.

After the researcher had reached an agreement with both forward-translators about the meaning of individual items, the three forward-translations were combined into one.

Optimally, a meeting would have been arranged between the researcher and forward-translators, but this was not possible since all three were living in different countries at the time. An online meeting was suggested by the researcher, but this was not possible due to scheduling conflicts. Instead, the researcher used his own judgment to combine the three versions. No difficulties relating to the meaning of individual items were encountered during the combination of the versions.

Back-Translation of the IBI. A form analogous to the forward-translation form was provided to the two back-translators, and the results were used to identify changes of meaning that might have occurred during the forward-translation. The type of feedback generated by the forward-translators was markedly different from the type of feedback from the back-translators. Whereas the forward-translators gave extensive feedback on the overall tenor and feel of the IBI but less feedback on individual items, the back-translators' feedback, in contrast, was much more clear-cut and focused on individual items rather than the tone. This may be due to the extensive editing done by the researcher and the forward-translators. Neither back-translator identified any culturally inappropriate items, but flagged several items as difficult and provided specific feedback. The results from back-translator 1 and back-translator 2 are found in Tables 4 and 5, respectively.

Table 4
Items Flagged by Back-Translator 1

Original English version	Back-translation	Preliminary Icelandic version	Comment
11. I avoid facing my problems (PA).*	I avoid confronting my problems.	Ég forðast að horfast í augu við vandamál mín.	I used “confronting”, but “facing” might be more literal.
12. A person won’t stay angry or blue long, unless he/she keeps himself/herself that way. (EI)*	You never feel angry or down for a long time unless you keep yourself feeling that way.	Maður er almennt ekki reiður eða einmana mjög lengi nema maður haldi sér þannig sjálfum.	I used “down” for “daufur í dálkinn”, but alternatives might include “blue” or “mildly depressed”.
23. I hate to fail at anything. (NFA)*	I can’t stand it when I fail to do something right.	Ég þoli alls ekki að mér mistakist eitthvað.	I felt like “fail to do something right” conveyed the meaning of “mistekst eitthvað” better than “make a mistake” (“verða á mistök”).
27. If something is necessary, I do it even if it is unpleasant. (PA)*	If something is necessary, I do it, even though it is inconvenient.	Ef ég þarf að gera eitthvað, þá geri ég það, jafnvel þó það sé óþægilegt.	“Óþægilegt” can be either “inconvenient” or “uncomfortable”.
33. More people should face up to the unpleasantness of life. (R)*	More people should face how unpleasant life can be.	Fleira fólk þarf að horfast í augu við að lífið er stundum erfið.	I feel like my translation is awkward, but couldn’t think of a better way to convey the meaning.
40. I often spend more time trying to think of ways of getting out of things than it would take me to do them. (PA)*	When I need to do something that’s unpleasant, I often spend more time thinking about how to get out of it than I would spend actually doing it.	Ég eyði oft meiri tíma í að hugsa um hvernig ég kemst hjá því að gera eitthvað en það tæki mig að gera það.	#40 and #44: Although “leiðinlegt” is used in both, I felt like “unpleasant” was a better match in the context of #40 and “boring” in the context of #44, but I may be wrong.
44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant. (PA)*	You should resist doing work that’s boring for as long as you possibly can, no matter how important it is.	Maður ætti að streitast á móti því að vinna leiðinleg og óþægileg verk, alveg sama hversu nauðsynleg þau eru.	#40 and #44: Although “leiðinlegt” is used in both, I felt like “unpleasant” was a better match in the context of #40 and “boring” in the context of #44, but I may be wrong.
46. People make their own hell within themselves. (EI)*	We each create our own fortune or misfortune.	Eymd og vanlíðan fólks er undir fólki sjálfu komið.	Wasn’t sure about this one, because it was based on an Icelandic saying, and I couldn’t think of a comparable English saying, although I am sure one exists.

* Subscale abbreviations: PA= Problem Avoidance, EI= Emotional Irresponsibility, NFA= Need For Approval, R= Rigidity.

Back-translator 2 identified no inappropriate items, but gave comments on the meaning and appropriateness of the following items (Table 5).

Table 5
Items Flagged by Back-Translator 2

English version	Back-translation	Preliminary Icelandic version	Comment
4. People should observe moral laws more strictly than they do. (R)*	People should be more moral.	Fólk ætti að virða almenn siðaboð og siðferðisleg gildi betur en það gerir.	Siðaboð is a somewhat archaic word.
12. A person won't stay angry or blue long, unless he/she keeps himself/herself that way. (EI)*	If you are angry or sad for long periods of time, you are doing it to yourself.	Maður er almennt ekki reiður eða einmana mjög lengi nema maður haldi sér þannig sjálfum.	Item 12 contains an expression that is difficult to translate: "nema maður haldi sér þannig sjálfum".
14. Those who do wrong deserve to be blamed. (R)*	People who do wrong deserve to be judged.	Fólk sem gerir illt á skilið að vera sakfellt.	"Sakfellt" in this item could be possibly translated as "judged", "convicted", and "condemned".
18. A large number of people are guilty of bad sexual conduct. (R)	A lot of people engage in amoral sexual behavior.	Margt fólk er sekt um slæma kynhegðun.	It is unclear what "slæm kynhegðun"[bad sexual conduct, sic] refers to. Is it dangerous to your health? Amoral? Harmful to others?
34. Helping others is the very basis of life. (R)	Helping others is the foundation for everything that is important in life.	Að hjálpa öðrum er undirstaða alls.	"Undirstaða alls" sounds strange in Icelandic. What is the "everything" that is being referred to?
21. I often worry about how people approve of and accept me. (NFA)	I often worry about people not respecting me or accepting me for who I am.	Ég hef oft áhyggjur af því hvort fólki líki vel við mig og hvaða álit það hefur á mér.	Viðurkenning (21, 37, 48) can possibly be translated as "respect" and "acknowledgement".
37. It is important to me that others approve of me. (NFA)	It is important to me that others acknowledge me.	Það er mér mikilvægt að aðrir viðurkenni mig.	Viðurkenning (21, 37, 48) can possibly be translated as "respect" and "acknowledgement".
48. Although I like approval, it's not a real need for me. (NFA)	Getting the respect of others is a good thing, but it is not something that I absolutely need to have.	Þótt mér þyki gott að falla öðrum í geð, þá er það mér samt ekki nauðsynlegt.	Viðurkenning (21, 37, 48) can possibly be translated as "respect" and "acknowledgement".

*Subscale abbreviations: R=Rigidity, EI= Emotional Irresponsibility, NFA= Need for approval

Only one item raised flags by both back-translators, albeit for different reasons. The item was item 12: "A person won't stay angry or blue long, unless he/she keeps himself/herself that way." On that item, back-translator 1 had difficulties translating the meaning of the term "blue", whereas back-translator 2 reported difficulties in translating "keeping themselves that way", since the word "keep" translates more to the keeping of material possessions rather than maintaining a mental state. The two back-translated versions of the IBI were judged to be relatively similar, the changes shown in Table 6 highlighting the major differences or disagreements. The researcher exercised final

judgment over the back-translated version, but the back-translators did not provide feedback on the final version of the IBI.

Cross-Cultural Equivalence of the IBI. Three cultural reviewers analyzed the Icelandic IBI and checked for relevance, meaning, and tone. As before, a form was given to the reviewers to facilitate comparisons (See Appendix E). Since most cultural commentators and translators tended to comment on the same items, and one commentator did not identify any significant problems with cultural equivalence, comments are combined in Table 6.

Table 6
Items Flagged by Cultural Evaluators

Item in English	Item in Icelandic	Comment	Modification attempted
4. People should observe moral laws more strictly than they do.	Fólk ætti að virða almenn siðaboð og siðferðisleg gildi betur en það gerir.	Only religious fundamentalists would think of moral values as "laws."	Change emphasis to "common ethics" and "values" over "laws."
20. It is sinful to doubt the Bible.	Það er syndsamlegt að efast um það sem stendur í Biblíunni.	Will not help identify Irrational Beliefs among non-Christians.	None, since bias would be expected to remain consistent between cultures.
46. People make their own hell within themselves.	Eymd og vanlíðan fólks er undir fólki sjálfu komið.	Reference to Hell as a place will not make sense in this context; "Hell" too dramatic.	Icelandic term reworded to the effect of: "People create their own misery and malaise."
39. It is realistic to expect that there should be no incompatibility in marriage.	Það er raunhæft að búast við hnökralausu hjónabandi.	Double negatives confusing; "incompatibility" appropriate for machines and technology, not for humans.	"It is realistic to expect a problem-free marriage."
14. Those who do wrong deserve to be blamed.	Fólk sem gerir illt á allt vont skilið.	Translation possibly stronger: "...deserve all bad things".	None, since "blame" in English is general.

As was the case during forward-translation, the most difficult items in the final cultural equivalence check were ones dealing with moral reasoning and ethics. Terms such as "moral laws," "blame," and "wrong," can all take on different meanings in Icelandic. One cultural reviewer found the IBI to have a strong and overt Christian overtone despite previous attempts to tone it down after similar comments were raised

during the forward-translation and back-translation rounds. Cultural reviewers were given the same comment form as the forward-translators and back-translators. To the question: "Did any items raise flags because of strange/archaic/culturally inappropriate wording or that you would not expect to see in a psychological test?", a reviewer found item 14, "Those who do wrong deserve to be blamed" difficult to edit due to multiple meanings. When translated to Icelandic, the word "blame" means "sök," which can mean "guilt" in both a legal sense or an abstract ethical sense. However, the word for "blame" in Icelandic is "ásökun," which means "accusation" in English, but is more often used as a verb than a noun. An email conversation with the reviewer revealed that the Icelandic reviewer had a more specific meaning in mind when reading the IBI, but the researcher decided to widen the focus on item 14 from specific guilt to literally translate to "those who do wrong deserve all bad things," to keep the tone of the IBI consistent to generalized ideas.

The U.S. and Icelandic versions of the IBI were somewhat different in length; the U.S. version is 546 words and 2349 characters (spaces not counted), but the Icelandic version is 664 words and 3009 characters. The full U.S. and Icelandic versions of the IBI are shown in Appendices A and B, respectively.

Phase 2: Cognitive Interviews and Cultural Validation of the IBI

A major objective of the study was to establish the extent of cultural equivalence, or lack of it, between the two versions of the IBI. To accomplish this, qualitative interview strategies were used to complement the psychometric analysis. In many ways, qualitative interviews yield data similar to those acquired during the translation phase, but since the translation is on the developer end of the study, and the cognitive interview

data come from the test-taker end, results from the cognitive interviews will be presented in the current phase, separately from the translator data.

As part of the data collection for the current study, one-on-one cognitive interviews were done in both countries to provide insights that might prove helpful in interpreting the psychometric results. To gain insight into the items, a group discussion was also conducted on a separate U.S. sample of 21 undergraduates at an urban research university. Since the objective of the study was not to modify the U.S. version of the IBI, the primary purpose of the qualitative judgment procedures was to aid interpretation. In Iceland, however, time constraints did not allow changes to be made to the IBI from the cognitive interviews. Instead, results from the Icelandic cognitive interviews were used to aid interpretation after the quantitative data had been collected. In both countries, the cognitive interviews were done using a combination of two techniques. First, participants were asked to respond to the IBI and verbalize their thoughts aloud as they responded to each item. Meanwhile, the researcher made notes on participants' comments or other visible reactions observed during the administration. This procedure was followed by the researcher asking probing questions about specific items, based on either a remark or an otherwise notable reaction such as hesitation, a giggle, or a facial expression. Since participants did not necessarily all respond to the same items, it was found more parsimonious to group the content of the comments by item responded to, rather than by participant. In the current chapter, results from the cognitive interviews for each country are presented first, followed by a the quantitative results from the psychometric analysis.

Cognitive Interviews and Discussion Group – U.S. A group interview was carried out in a classroom with 21 students. The group discussion focused on the English-version

of the IBI. The researcher made notes on each comment and asked the group questions if needed. A summary of the comments is presented in Table 7.

Table 7
*Items Identified as Problematic by U.S. Discussion Group**

Item # and description	Researcher notes on discussion group comment
3. Certain people are bad or wicked and should be severely punished for their sins.	<ul style="list-style-type: none"> • "Wicked" sounds a bit "off". • Strange archaic wording, "severely punished".
8. I am fairly easygoing about life	<ul style="list-style-type: none"> • Will anyone say no to this?
10. I hardly ever think of such things as death or atomic war.	<ul style="list-style-type: none"> • "Atomic war" not something to fear anymore. • Terrorism more appropriate threat. • Archaic wording.
14. Those who do wrong deserve to be blamed.	<ul style="list-style-type: none"> • Puzzlement over wording. "Punished" would have been clearer.
25. I shrink from facing a crisis or difficulty.	<ul style="list-style-type: none"> • "Shrink from" is unclear wording.
26. I feel little anxiety over unexpected danger or future events.	<ul style="list-style-type: none"> • Stiff wording • Unclear, needed to re-read a few times to understand what "little" meant.
28. Frustrations upset me.	<ul style="list-style-type: none"> • Stating the obvious? • No apparent semantic difference between being "upset" and "frustrated". • The double negatives are confusing
39. It is realistic to expect that there should be no incompatibility in marriage	<ul style="list-style-type: none"> • The double negatives are confusing
41. Immorality should be strongly punished.	<ul style="list-style-type: none"> • Is this legally possible? • Who is to decide what immorality is? God? Society? The courts? Etc.
43. What others think of you is most important	<ul style="list-style-type: none"> • The most important thing in life, or just very important? Unclear wording.
44. One should rebel against doing unpleasant things, however, necessary, if doing them is unpleasant.	<ul style="list-style-type: none"> • "Rebel" against whom? Strange choice of words.
46. People make their own hell within themselves.	<ul style="list-style-type: none"> • Religious views and lack thereof induced different interpretations, literal vs. metaphorical.
49. People who are miserable have usually made themselves that way.	<ul style="list-style-type: none"> • Does "miserable" refer to a state of mind or unfortunate circumstances?

*Responses are grouped and summarized into bullet points by researcher.

In the group discussion, comments relating to wording were the most common. The cognitive interviews were carried out in a one-on-one setting. After the think-aloud procedure, the researcher used probes to elicit additional responses from participants. A summary of the results is in Table 8.

Table 8
*Items Identified as Problematic in U.S. Cognitive Interviews**

Item # and description	Researcher's notes
3. Certain people are bad or wicked and should be severely punished for their sins.	<ul style="list-style-type: none"> • More for their crimes than their sins. • What is bad? • How severe is the punishment?
4. People should obey moral laws more strictly than they do	<ul style="list-style-type: none"> • Aren't all laws supposed to be moral? • Which moral laws?
10. I hardly ever think of such things as death or atomic war.	<ul style="list-style-type: none"> • "I don't think of atomic war at all". • Is atomic war a threat anymore? • "Was this written during the Cold War?" • Would "nuclear" be better? Terrorism?
17. Nothing is upsetting in itself - only in the way you interpret it.	<ul style="list-style-type: none"> • But sometimes bad things do happen, accidents, loss, etc.
18. A large number of people are guilty of bad sexual conduct.	<ul style="list-style-type: none"> • Not sure what "bad" means. • "Guilty" meaning by law?
19. I often get excited or upset when things go wrong.	<ul style="list-style-type: none"> • Student came straight to the interview after a big test and claimed this was affecting her response to the item.
21. I often worry about how people approve of and accept me.	<ul style="list-style-type: none"> • "Only people I care about"
25. I shrink from facing a crisis or a difficulty.	<ul style="list-style-type: none"> • "Shrink" understood as "frustrated" • "Shrink" understood as "feeling powerless". • "Shrink" not understood.
26. I feel little anxiety over unexpected danger or future events.	<ul style="list-style-type: none"> • Student reported anxiety over graduation at time of interview.
29. One should blame oneself severely for all mistakes and wrongdoings.	<ul style="list-style-type: none"> • "Severely" too strong? • "Severely" unclear. Substitute for "harsh"?

*Bullets denote responses from multiple respondents

Table continues on next page

Table 8 Continued
*Items Identified as Problematic in U.S. Cognitive Interviews**

Item # and description	Researcher's notes
32. I get terribly upset when things are not the way I like them to be.	<ul style="list-style-type: none"> • "Terribly" too strong?
33. More people should face up to the unpleasantness of life.	<ul style="list-style-type: none"> • Unclear whether "unpleasantness of life" is implied as a generalization, or as a reference to the odd unpleasant episode.
36. It is difficult for me to do unpleasant chores.	<ul style="list-style-type: none"> • Difficult to do vs. difficult to get around doing?
38. Too many evil persons escape the punishment they deserve.	<ul style="list-style-type: none"> • Notes tedium by the recurrence of the item on the list.
38. Too many evil persons escape the punishment they deserve.	<ul style="list-style-type: none"> • What is meant by "evil"? • Punishment by whom, karma or law? "Evil" is ambiguous.
39. It is realistic to expect that there should be no incompatibility in marriage.	<ul style="list-style-type: none"> • Very complicated item. Double negative confusing. • Confused by double negative.
43. What others think of you is most important.	<ul style="list-style-type: none"> • Notes similarity of item to several other items.
44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant.	<ul style="list-style-type: none"> • "Rebel or resist"? • Nonspecific confusion about the item. • Comments on "strange wording".
45. I can't stand to take chances.	<ul style="list-style-type: none"> • Feels "I can't stand" is too strong.
46. People make their own hell within themselves.	<ul style="list-style-type: none"> • Interprets "hell" as "bad conscience", wonders whether this is the correct interpretation.
47. I dislike responsibility	<ul style="list-style-type: none"> • Interprets responsibility as "being responsible", wonders how anyone could dislike that. • "Responsibility makes me feel good".
50. I have considerable concern with what people are feeling about me.	<ul style="list-style-type: none"> • Admits trying to think of what she answered previously to similar questions [response set induced]

*Bullets denote responses from multiple respondents.

The comments from U.S. participants in both the group discussions and cognitive interviews centered mostly on the apparently archaic and less-than-fluid wording of the IBI, but many centered on the relevance of the items themselves. Value-laden words such

as "evil", "bad", "blame", "terribly", and "severely", which occurred frequently, tended to be seen as strange or inappropriate to the context of each item. For example, "atomic war", as referred to in the item, 10, "I hardly ever think of such things as death or atomic war", caused pause by all participants in the cognitive interviews as well as several group discussion participants. The fear of nuclear war did not appear to be salient among participants, but other events in the public discourse at the time were suggested instead, including but not limited to terrorism, the rise of hostile superpowers, home invasions, and campus gun violence. A couple of students noted that some items seemed to appear over and over again with little change in wording. Some students were surprised by the apparent religious content of some items.

Cognitive Interviews - Iceland. It was not possible to conduct group discussion in Iceland due to time limitations on data collection in the country. However, cognitive interviews were conducted, and participants in the actual administration of the IBI on the Iceland side were informally interviewed after each session whenever possible. For more detailed comments, see Table 9.

Table 9
*Items Identified as Problematic in Icelandic Cognitive Interviews**

Item # and description	Icelandic wording	Researcher's notes
4. People should obey moral laws more strictly than they do.	Fólk ætti að virða almenn siðaboð og siðferðisleg gildi betur en það gerir.	<ul style="list-style-type: none"> Some confusion about whether this applies to laws of the land or ethics in the more abstract sense.
12. A person won't stay angry or blue long, unless he/she keeps himself/herself that way.	Fólk er yfirleitt ekki reitt eða einmana mjög lengi nema það velti sér upp úr eigin eynd.	<ul style="list-style-type: none"> "Sometimes things happen to people that they can't control".
13. I usually try to avoid chores that I dislike doing.	Ég reyni yfirleitt að koma mér undan skylduverkum sem mér þykja leiðinleg.	<ul style="list-style-type: none"> "Isn't that a moot point?" (chore translates more closely to "work of duty" rather than something tedious.)
15. If one wants to, one can be happy under almost any circumstances.	Ef viljinn er fyrir hendi er hægt að vera hamingjusamur við nánast hvaða kringumstæður sem er.	<ul style="list-style-type: none"> Some circumstances are inherently unpleasant (circumstances being understood very specifically).
18. A large number of people are guilty of bad sexual conduct.	Margt fólk hegðar sér illa í kynferðismálum.	<ul style="list-style-type: none"> Not sure what "bad" means. "Guilty" meaning by law?
21. I often worry about how people approve of an accept me.	Ég hef oft áhyggjur af því hvort fólki líki vel við mig og hvaða álit það hefur á mér.	<ul style="list-style-type: none"> Being approved of (Icelandic: Liked), and being accepted (what they think of me) are not necessarily the same.
24. The fear of punishment helps people to be good.	Hræðslan við refsingu heldur fólki heiðarlegu.	<ul style="list-style-type: none"> "If this were the case, everybody would be honest" - took it very literally and generally.
29. One should blame oneself severely for all mistakes and wrongdoings.	Maður á að taka harkalega í lurginn á sjálfum sér fyrir eigin mistök og misgjörðir.	<ul style="list-style-type: none"> Noted that this question appeared before.
32. I get terribly upset when things are not the way I like them to be.	Ég verð afskaplega æst(ur) og mér líður ömurlega þegar hlutirnir eru ekki eins og ég vil hafa þá.	<ul style="list-style-type: none"> "Are these all the same questions over again?" "I think I just answered that".
35. There is a right way to do everything.	Það er til rétt aðferð við að gera hvað sem er.	<ul style="list-style-type: none"> Had to read the item several times to understand. "I understand it but it sounds weird". "Do you mean ONE right way to do everything?" "Again?" This sounds exactly like an earlier item.
36. It is difficult for me to do unpleasant chores.	Ég á erfitt með að gera leiðinleg skylduverk.	<ul style="list-style-type: none"> This sounds exactly like an earlier item.
37. It is important to me that others approve of me.	Það er mér mikilvægt að aðrir viðurkenni mig.	<ul style="list-style-type: none"> "Now this list is just repeating itself" (loses interest).
38. Too many evil persons escape the punishment they deserve.	Of mörg illmenni sleppa við refsinguna sem þau verðskulda.	<ul style="list-style-type: none"> "Are you testing Icelandic attitudes to the banker crooks?"
46. People make their own hell within themselves.	Eynd og vanlíðan fólks er undir fólki sjálfu komið.	<ul style="list-style-type: none"> Sounds judgmental.
49. People who are miserable have usually made themselves that way.	Fólki sem líður illa hefur yfirleitt komið sjálfu sér í þá stöðu.	<ul style="list-style-type: none"> This sounds harsh. Notes similarity with item 46 and others on IBI.

*Bullets denote multiple respondents

Cognitive Interviews - Overview. Several results emerged from the cognitive interviewing stage. One very clear result was that all the Icelanders interviewed perceived an item or even a set of items as extremely similar to ones they had responded to before. This was also observed after the formal classroom administration of the IBI in Iceland, in conversations or debriefings with participants. While some of the U.S. cognitive interviewees and discussion group participants also found the IBI to be repetitive, this impression was noticeable in the U.S. but was perhaps the most salient finding from the cognitive interviews in Iceland. A shared confusion about the same issue among the Icelandic and U.S. interviewees appeared only on item 4, "People should obey moral laws more strictly than they do". Participants in both countries found the item ambiguous and wondered whether "moral laws" referred to laws of the land or abstract ethics (Tables 8 and 9).

Phase 3: A Psychometric Analysis of the IBI

Descriptive statistics and item-to-total correlations were obtained using SPSS statistical software, version 18 for the 50 IBI items for each country. Item descriptives and item-to-total correlations are shown in Table 11. A few questions on the list needed to be reverse-coded due to wording of the items, which was done in SPSS 18. Since the data were nested in classrooms, the data were tested for violation of the independence assumption, that is, whether the individuals within the classes were more homogeneous than if they had been sampled at random. The assumption of independence was evaluated by calculating the intraclass correlation coefficient (ICC) for each item by country. The ICC represents the variation between classes divided by the total variability. The independence assumption was not found to be violated in either sample. In the Icelandic

sample, the range of ICCs was .001 to .123, with a median of .016 and a mode of .026. In the U.S. sample, the range of ICCs was .000 to .045, with a median of .018 and a mode of .022.

For each of the two countries involved, confirmatory factor analysis (CFA) was performed to evaluate the five-factor model of the Irrational Beliefs Inventory, using Mplus statistical software, version 4.2 (Muthén & Muthén, 1998-2007). Analyses were based on the variance-covariance matrix of the 50 items, and maximum likelihood estimation was used to estimate model parameters. The fit indices used were the chi-square test (χ^2), the comparative fit index (CFI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Tests of statistical significance were conducted at the .05 level. The initial CFAs were conducted separately for each country, followed by an invariance test between the two countries. Full information maximum likelihood estimation (ML) using the EM algorithm was used to handle missing data. Several items in both countries had no missing cases, but the item with the highest missing values was item 39 in the U.S. sample, with 3.0% of the responses missing. In Iceland, the highest proportion of missing data was 1.11% on item 26. For a further breakdown on missing data, see Table 10.

Table 10
Percentage of Missing Data across Countries, Grouped by Item

Item	% missing	
	U.S.	Iceland
1. If I can't keep something from happening, I don't worry about it	0.36	0.56
2. I worry a lot about certain things in the future	0.13	0.00
3. Certain people are bad or wicked and should be severely punished for their sins	0.36	0.14
4. People should observe moral laws more strictly than they do	0.12	0.42
5. I want everyone to like me	0.24	0.14
6. I often can't get my mind off some concern	0.36	0.28
7. I tend to become terribly upset when things are not the way I would like them to be.	0.00	0.28
8. I am fairly easygoing about life	0.00	0.14
9. Punishing oneself for all errors will prevent future mistakes	0.48	0.56
10. I hardly ever think of such things as death or atomic war	0.60	0.14
11. I avoid facing my problems	0.24	0.42
12. People don't stay angry or blue long, unless they keep themselves that way	0.48	0.14
13. I usually try to avoid chores which I dislike doing	0.24	0.28
14. Those who do wrong deserve to be blamed	0.73	0.42
15. If one wants to, one can be happy under almost any circumstances	0.48	0.14
16. I tend to worry about possible accidents and disasters	0.48	0.00
17. Nothing is upsetting in itself - only in the way you interpret it	0.73	0.28
18. A large number of people are guilty of bad sexual conduct	1.09	0.14
19. I often get excited or upset when things go wrong	0.48	0.0
20. It is sinful to doubt the Bible	1.45	0.00
21. I often worry about how people approve of and accept me	0.60	0.00
22. Sometimes I can't get a fear off my mind	0.12	0.00
23. I hate to fail at anything	0.24	0.00
24. The fear of punishment helps people to be good	0.36	0.56
25. I shrink from facing a crisis or difficulty	0.69	0.69
26. I feel little anxiety over unexpected danger or future events	0.85	1.11
27. If something is necessary, I do it even if it is unpleasant	1.09	0.56
28. Frustrations upset me	1.57	0.56
29. One should blame oneself severely for all mistakes and wrongdoings	1.45	0.14
30. People are disturbed not by situations but by the view they take of them	1.93	0.42
31. I usually put off important decisions	0.69	0.28
32. I get terribly upset and miserable when things are not the way I like them to be	1.33	0.56
33. More people should face up to the unpleasantness of life	2.42	0.14
34. Helping others is the very basis of life	0.85	0.28
35. There is a right way to do everything	1.45	0.56
36. It is difficult for me to do unpleasant chores	1.33	0.56
37. It is important to me that others approve of me	1.21	0.42
38. Too many evil persons escape the punishment they deserve	1.69	0.69
39. It is realistic to expect that there should be no incompatibility in marriage	3.02	0.56
40. I often spend more time trying to think of ways of getting out of things than it would take me to do them	1.45	0.69
41. Immorality should be strongly punished.	2.42	0.69
42. There is never any reason to remain sorrowful for very long	1.69	0.83
43. What others think of you is most important	1.33	0.83
44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant	2.06	0.97
45. I can't stand to take chances	1.09	0.56
46. People make their own hell within themselves	2.30	0.11
47. I dislike responsibility	0.97	0.69
48. Although I like approval, it's not a real need for me	1.33	0.69
49. People who are miserable have usually made themselves that way	1.09	0.56
50. I have considerable concern with what people are feeling about me	1.09	0.28

U.S. Sample Statistics. Items on the IBI are 50 statements with which respondents are asked to indicate a range of reactions, from strong disagreement to strong agreement.

Responses are arranged on a 5-point Likert scale, ranging from 1 = Strongly Disagree,

2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree. Sample statistics for the U.S. sample are shown in Table 11. The neutral option is defined as part of the opinion continuum on the IBI, so it was factored in the mean. Skewness and kurtosis statistics were used to gauge the normality of the distributions for the items, and item-to-total correlations were used to check the items' consistency with their corresponding subscales. Item 47 had a skewness value of +1.05 and was the only item skewed beyond -1 or 1. Ten items were kurtotic (4,10, 16, 20, 22, 26, 27, 28, 45, 47), out of which four items were kurtotic beyond -1 and 1 on both Icelandic and U.S. versions (4, 10, 26, and 27.) The highest kurtosis values in the U.S. sample were at -1.09 and +2.38 for items 20 and 27, respectively.

Table 11
Item Descriptives and Item-to-Total Correlations for the U.S. Sample (n = 827)*

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Worrying	1. If I can't keep something from happening, I don't worry about it.**	3.21	1.11	-0.34	-0.94	.41
	2. I worry a lot about certain things in the future.	3.78	0.96	-0.82	0.36	.49
	6. I often can't get my mind off some concern.	3.63	0.95	-0.70	0.06	.55
	7. I tend to become terribly upset when things are not the way I like them to be.	3.03	1.07	0.11	-0.94	.57
	8. I am fairly easygoing about life.**	2.16	0.90	0.65	0.08	.38
	10. I hardly ever think of such things as death or atomic war.**	3.00	1.17	-0.03	-1.03	.20
	16. I tend to worry about possible accidents and disasters.	3.00	1.05	-0.11	-1.00	.48
	19. I often get excited or upset when things go wrong.	3.32	0.95	-0.45	-0.53	.51
	22. Sometimes I can't get a fear off my mind.	3.00	1.09	-0.13	-1.04	.53
	26. I feel little anxiety over unexpected danger or future events.**	2.92	1.00	0.18	-1.00	-.02
28. Frustrations upset me.	3.79	0.76	-0.85	1.29	.44	
32. I get terribly upset and miserable when things are not the way I like them to be.	2.81	1.04	0.29	-0.74	.52	

** Rationally worded items that were reverse-coded

Table continues on next page

Table 11 Continued

*Item Descriptives and Item-to-Total Correlations for the U.S. Sample (n = 827)**

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Rigidity	3. Certain people are bad or wicked and should be severely punished for their sins.	2.91	1.07	-0.06	-0.60	.44
	4. People should observe moral laws more strictly than they do.	3.70	0.84	-0.82	1.05	.46
	9. Punishing oneself for all errors will prevent future mistakes.	2.13	0.89	0.80	0.48	.37
	14. Those who do wrong deserve to be blamed.	3.49	0.87	-0.52	0.00	.37
	18. A large number of people are guilty of bad sexual conduct.	3.13	0.98	-0.19	-0.39	.40
	20. It is sinful to doubt the Bible.	2.66	1.32	0.22	-1.09	.39
	24. The fear of punishment helps people to be good.	3.23	1.04	-0.51	-0.52	.36
	29. One should blame oneself severely for all mistakes and wrongdoings.	2.18	0.90	0.75	0.52	.39
	33. More people should face up to the unpleasantness of life.**	3.13	0.92	-0.16	-0.40	.23
	34. Helping others is the very basis of life.	3.98	0.78	-0.64	0.66	.09
	35. There is a right way to do everything.	2.80	1.14	0.25	-0.78	.40
	38. Too many evil persons escape the punishment they deserve.	3.28	1.00	-0.37	-0.33	.49
	39. It is realistic to expect that there should be no incompatibility in marriage.	2.26	0.94	0.45	-0.18	.24
41. Immorality should be strongly punished.	2.77	0.91	-0.05	-0.04	.52	

** Denotes items that were worded in a rational direction and were reverse-coded
 Table continues on next page

Table 11 Continued
*Item Descriptives and Item-to-Total Correlations for the U.S. Sample (n = 827)**

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Problem avoidance	11. I avoid facing my problems.	2.18	0.91	0.77	0.25	.49
	13. I usually avoid chores which I dislike doing.	3.17	1.07	-0.21	-1.02	.38
	25. I shrink from facing a crisis or difficulty.	2.49	0.87	0.38	-0.28	.38
	27. If something is necessary, I do it even if it is unpleasant.**	2.09	0.70	0.99	2.38	.21
	31. I usually put off important decisions.	2.51	1.05	0.54	-0.49	.46
	36. It is difficult for me to do unpleasant chores.	2.77	0.98	0.24	-0.85	.48
	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them.	2.44	1.03	0.51	-0.51	.49
	44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant.	2.21	0.85	0.62	0.64	.24
	45. I can't stand to take chances.	2.12	0.81	0.87	1.10	.29
47. I dislike responsibility.	1.96	0.84	1.05	1.47	.42	
Need for approval	5. I want everyone to like me.	3.24	1.00	-0.22	-0.52	.61
	21. I often worry about how people approve of and accept me.	3.13	1.05	-0.17	-0.76	.72
	23. I hate to fail at anything.	3.92	0.97	-0.88	0.31	.21
	37. It is important to me that others approve of me.	3.12	1.04	-0.17	-0.74	.78
	43. What others think of you is most important.	2.20	0.96	0.62	-0.08	.50
	48. Although I like approval, it's not a real need for me.**	2.66	1.03	0.27	-0.82	.45
	50. I have considerable concern with what people are feeling about me.	2.90	1.03	-0.05	-0.81	.71
Emotional Irresponsibility	12. A person won't stay angry or blue long, unless he/he keeps himself/herself that way.**	2.38	0.98	0.66	-0.10	.48
	15. If one wants to, one can be happy under almost any circumstances.**	2.28	0.99	0.72	-0.03	.48
	17. Nothing is upsetting in itself - only in the way you interpret it.**	3.01	1.03	-0.19	-0.82	.45
	30. People are disturbed not by situations but by the view they take of them.**	2.65	0.89	0.30	-0.29	.49
	42. There is never any reason to remain sorrowful for very long.**	3.02	1.03	-0.10	-0.89	.47
	46. People make their own hell within themselves.**	2.72	0.98	0.31	-0.38	.46
	49. People who are miserable have usually made themselves that way.**	2.72	1.01	0.26	-0.67	.53

** Denotes items that were worded in a rational direction and were reverse-coded

Iceland Sample Statistics. Table 12 shows the descriptives and item-to-total correlations for the Icelandic sample. Three items had skewness values outside the range of -1 and +1 (20, 39, and 44); the highest negative skewness value was -0.99 for item 6, and the highest positive skewness value was +1.08 for item 20. Nine items showed kurtosis beyond -1 and +1 (1, 2, 4, 10, 21, 26, 27, 34, 44), with the highest values being -1.18 (item 1) and +1.75 (item 44), respectively.

Table 12
*Item Descriptives and Item-to-Total Correlations for the Icelandic Sample (n = 720)**

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Worrying	1. If I can't keep something from happening, I don't worry about it. **	3.09	1.11	-0.13	-1.18	.31
	2. I worry a lot about certain things in the future.	3.09	1.07	-0.21	-1.05	.60
	6. I often can't get my mind off some concern.	3.87	0.94	-0.99	0.62	.50
	7. I tend to become terribly upset when things are not the way I like them to be.	3.57	1.03	-0.60	-0.42	.53
	8. I am fairly easygoing about life.**	2.61	1.02	0.34	-0.72	.54
	10. I hardly ever think of such things as death or atomic war.**	2.87	1.17	-0.03	-1.11	.28
	16. I tend to worry about possible accidents and disasters.	2.60	1.15	0.33	-0.90	.47
	19. I often get excited or upset when things go wrong.	3.23	1.00	-0.30	-0.97	.57
	22. Sometimes I can't get a fear off my mind.	3.86	0.93	-0.92	0.47	.55
	26. I feel little anxiety over unexpected danger or future events.**	2.93	1.06	-0.03	-1.07	.50
	28. Frustrations upset me.	3.33	0.96	-0.44	-0.77	.52
32. I get terribly upset and miserable when things are not the way I like them to be.	2.72	0.98	0.31	-0.76	.51	

** Denotes items that were worded in a rational direction and were reverse-coded
Table continues on next page

Table 12 Continued
*Item Descriptives and Item-to-Total Correlations for the Icelandic Sample (n = 720)**

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Rigidity	3. Certain people are bad or wicked and should be severely punished for their sins.	3.34	1.08	-0.40	-0.65	.41
	4. People should observe moral laws more strictly than they do.	3.91	0.70	-0.64	1.24	.22
	9. Punishing oneself for all errors will prevent future mistakes.	1.96	0.85	0.90	0.82	.35
	14. Those who do wrong deserve to be blamed.	2.46	0.94	0.45	-0.42	.33
	18. A large number of people are guilty of bad sexual conduct.	3.72	0.85	-0.46	0.20	.20
	20. It is sinful to doubt the Bible.	1.74	0.96	1.08	0.40	.30
	24. The fear of punishment helps people to be good.	2.68	0.99	0.18	-0.96	.33
	29. One should blame oneself severely for all mistakes and wrongdoings.	2.32	0.94	0.47	-0.37	.35
	33. More people should face up to the unpleasantness of life.**	3.62	0.84	-0.54	0.39	.20
	34. Helping others is the very basis of life.	4.03	0.72	-0.91	2.11	.15
	35. There is a right way to do everything.	3.02	1.04	0.03	-0.85	.27
	38. Too many evil persons escape the punishment they deserve.	3.96	0.87	-0.72	0.53	.40
	39. It is realistic to expect that there should be no incompatibility in marriage.	2.00	1.00	1.01	0.49	.10
41. Immorality should be strongly punished.	3.17	0.93	-0.20	-0.19	.44	
Problem avoidance	11. I avoid facing my problems.	2.16	0.89	0.85	0.33	.46
	13. I usually avoid chores which I dislike doing.	2.45	1.00	0.66	-0.33	.54
	25. I shrink from facing a crisis or difficulty.	2.44	0.86	0.71	-0.06	.50
	27. If something is necessary. I do it even if it is unpleasant.**	2.26	0.78	0.99	1.12	.44
	31. I usually put off important decisions.	2.69	1.07	0.42	-0.86	.53
	36. It is difficult for me to do unpleasant chores.	2.63	0.99	0.48	-0.62	.54
	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them.	2.52	1.06	0.45	-0.65	.52
	44. One should rebel against doing unpleasant things. however necessary. if doing them is unpleasant.	1.73	0.71	1.02	1.75	.28
	45. I can't stand to take chances.	2.55	0.93	0.53	-0.43	.22
47. I dislike responsibility.	2.24	0.91	0.82	0.38	.47	

** Denotes items that were worded in a rational direction and were reverse-coded
 Table continues on next page

Table 12 Continued
*Item Descriptives and Item-to-Total Correlations for the Icelandic Sample (n = 720)**

Subscale	Item	Mean	SD	Skewness	Kurtosis	Corrected item-to-total correlations
Need for approval	5. I want everyone to like me.	3.63	0.96	-0.57	-0.20	.49
	21. I often worry about how people approve of and accept me.	3.19	1.11	-0.22	-1.03	.66
	23. I hate to fail at anything.	3.54	1.02	-0.42	-0.75	.36
	37. It is important to me that others approve of me.	3.55	0.91	-0.74	0.22	.61
	43. What others think of you is most important.	3.44	0.95	-0.48	-0.47	.76
	48. Although I like approval. it's not a real need for me.**	2.29	0.90	0.67	0.13	.40
	50. I have considerable concern with what people are feeling about me.	2.78	1.05	0.17	-0.92	.67
Emotional Irresponsibility	12. A person won't stay angry or blue long. unless he/he keeps himself/herself that way.**	2.63	0.98	0.40	-0.34	.37
	15. If one wants to. one can be happy under almost any circumstances.**	2.16	0.94	0.96	0.52	.39
	17. Nothing is upsetting in itself - only in the way you interpret it.**	3.01	0.10	-0.11	-0.96	.45
	30. People are disturbed not by situations but by the view they take of them.**	2.73	0.88	0.21	-0.41	.38
	42. There is never any reason to remain sorrowful for very long.**	3.25	1.06	-0.27	-0.86	.40
	46. People make their own hell within themselves.*	2.89	0.99	0.11	-0.87	.55
	49. People who are miserable have usually made themselves that way.**	3.44	0.95	-0.28	-0.59	.48

* Table 10 lists percentage of missing data by item

** Denotes items that were worded in a rational direction and were reverse-coded

Internal consistency for all the subscales in the questionnaires was tested using Cronbach's alpha (Table 13). Overall, the internal consistency ranged from .73 to .84 for the U.S. and ranged from .72 to .86 in Iceland. The item-to-total correlations tended to have a wider range in the U.S. overall, with the low end of the range slightly lower in the U.S. but the upper ranges similar in both countries across all subscales. Overall, the item-to-total correlations are acceptable but not overly high. Internal consistency was on the lower side considering the many items on the IBI.

Table 13
Internal Consistency of IBI Subscales (Cronbach's α)

Scale	# of items	U.S. Cronbach's α	U.S. n	U.S. Item-to-total correlation range	Iceland Cronbach's α	Iceland n	Iceland Item-to-total correlation range
Worrying	12	.78	791	-.02 to .57	.83	696	.28 to .60
Rigidity	14	.76	772	.09 to .52	.67	696	.10 to .44
Problem Avoidance	10	.73	788	.21 to .49	.78	693	.22 to .54
Need for Approval	7	.82	798	.21 to .78	.82	710	.41 to .67
Emotional Irresponsibility	7	.76	794	.45 to .53	.72	703	.37 to .52
All items	50	.84	713	-.03 to .55	.86	643	-.03 to .58

Confirmatory Factor Analysis of the IBI

In previous studies, the Dutch and U.S. versions of the IBI have yielded identical factor structures of the five subscales, and both have 50 items (Bridges & Sanderman, 2002; Koopmans et al., 1994). The five factors or conceptual categories that the IBI was designed to measure are worrying, rigidity, problem avoidance, need for approval, and emotional irresponsibility. For the current study, confirmatory factor analysis (CFA) was conducted to examine the five-factor structure of the IBI within both Iceland and the U.S.

A diagram of the basic original CFA model is displayed in Figure 1.

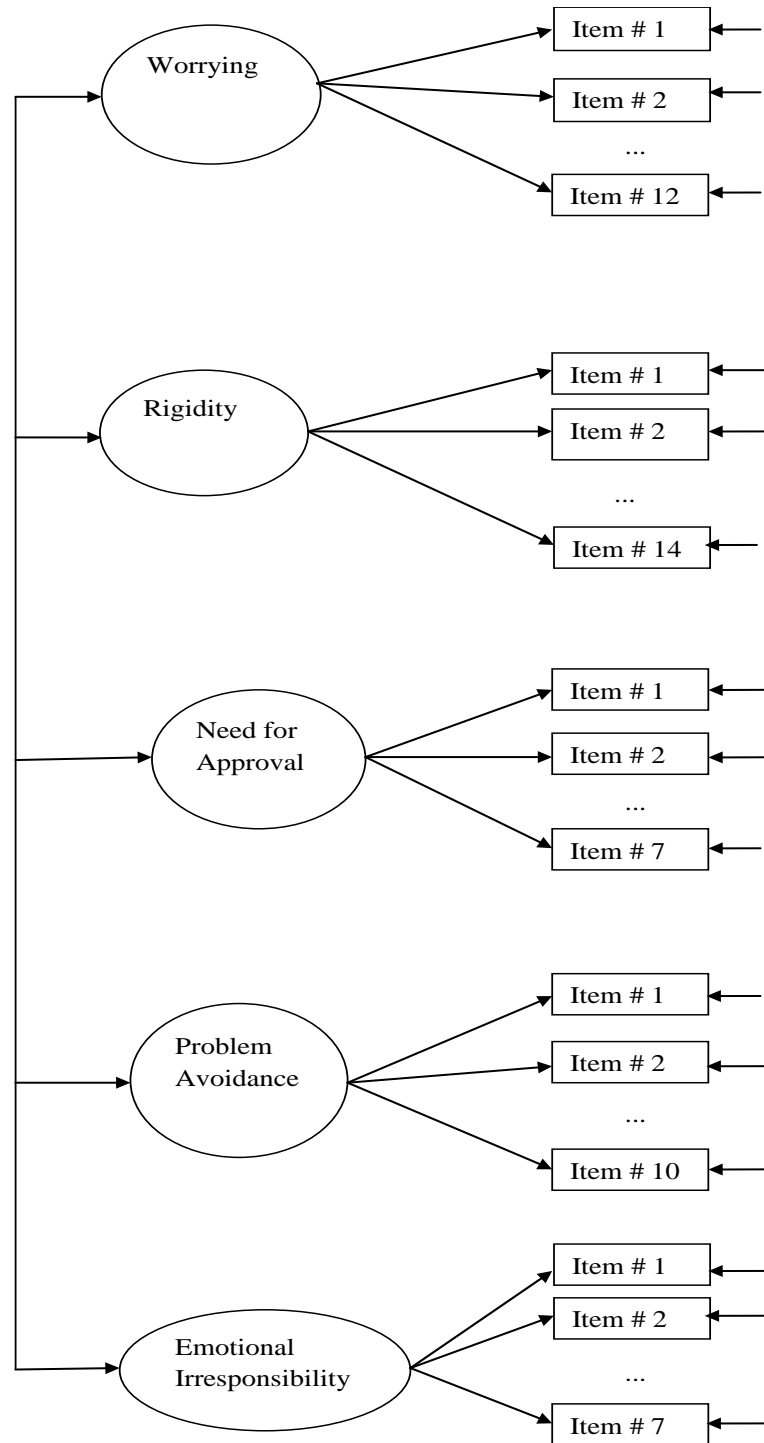


Figure 1. CFA model for the Icelandic adaptation of the IBI

Mplus statistical software, version 4 (Muthén & Muthén, 2005-2010) was used to test the five-factor model, and the input matrix was a covariance matrix. Fit of the five-

factor model was evaluated using a chi-square test of model fit, and other measures were the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR). Acceptable fit criteria were set at $CFI \geq .90$, $RMSEA < .08$, and $SRMR < .08$. Factor loadings, correlations between the five factors/subscales of the IBI, and modification indices indicating sources of misfit (i.e., changes in the model's chi-square) were also examined to get an idea how the different versions fit for each country (Table 14). The critical value of the chi-square change for one degree of freedom was 10.83 ($p < .001$), so only modification indices at or above 10.83 were displayed when each model was run.

Model Fit. A CFA was run on the Icelandic and U.S. samples individually to determine how well the data fit the original five-factor structure of the IBI. In addition, specific potential sources of misfit were examined. The general findings from the two CFAs are presented first, with more detailed findings on sources of misfit to follow in subsequent sections. Test of model fit results compared across countries are displayed in Table 14. In both Icelandic and U.S. samples, the chi-square and CFI suggested poor fit, while the RMSEA and SRMR suggested acceptable fit. The Icelandic model suggested a slightly better fit on the chi-square but was slightly worse on the other indices. It should be noted that there is considerable controversy surrounding interpretations of individual fit indices, and it is not uncommon that fit indices conflict, due to their different statistical makeup (Brown, 2006). The chi-square and SRMR are so-called absolute indices because they test a model fit regardless of other potential adjustments to the model (Brown, 2006). In addition, the CFI tests the fit of a model against the worst possible "null" model, whereas the RMSEA tests model fit without a null model assumption, which in

turn means that poor loadings in the model can bring down the RMSEA and the chi-square (Brown, 2006; Rigdon, 1996). The RMSEA is called a parsimony correction fit index, because it takes into account the complexity of the model, and the assumption is that it tests whether the model fits reasonably well in the population (Brown, 2006; Hu & Bentler, 1999). Hu and Bentler (1999) recommend that several fit indices be used to supplement the commonly used chi-square, as the chi-square is highly sensitive to sample size. All that being said, model evaluation should always be holistic and guided by theory rather than a unilaterally rigid emphasis on fit indices alone (Hu & Bentler, 1999; Rigdon, 1996). Discord of fit indices within each country notwithstanding, they were remarkably consistent across the countries, and the RMSEA and SRMR were near-identical (Table 14).

Modification indices are estimations of how much the chi-square will change if a parameter that has been fixed (e.g., 0), is freely estimated; thus, modification indices are potential identifiers of items that might contribute to a model's misfit (Brown, 2006). In the present models, two potential sources of misfit were identified. First, all of the errors associated with the items were assumed to be uncorrelated, and thus the presence of correlated errors between pairs of items would represent a source of misfit. If freeing the fixed parameter and estimating the correlation between the errors produced a statistically significant improvement in fit (as measured by the change in chi-square relative to the loss of a degree of freedom), this would provide evidence of a source of model misfit. A second source of potential misfit in the model can occur when an item has cross-loadings on one or more factors beyond the factor that the item was designed to represent. For both of these situations, the higher the correlated error and/or the second loading, the

worse the fit of the model, and the higher the modification index for that item will be (Brown, 2006).

Validity of a test item on an instrument such as the IBI comes from many sources, including but not limited to the item's conceptualization, to clarity of wording and the cultural appropriateness (Brown, 2006). Therefore, modification indices were used in the present study as flags to help identify potential items that might need improvement, but interpretations in the present study were focused just as much on the qualitative judgment strategies used to gain insight into the IBI's performance.

Table 14
Confirmatory Factor Analysis for the IBI, Icelandic and U.S. Samples

Fit Measure	Iceland (<i>n</i> = 720)	U.S. (<i>n</i> = 827)
χ^2	3177.818	3612.321
df	1165	1165
CFI	.76	.76
RMSEA	.05	.05
SRMR	.06	.06
Largest modification index	101.95	121.68
Description of largest modification index	Cross-loading of item 23 on Worrying factor	Correlated error between items 30 and 17

Generally speaking, the CFA of the U.S. sample in the present study did not yield a good fit with the five-factor model proposed in previous studies (Bridges & Sanderman, 2002; Koopmans et al., 1994). In the current U.S. model, a total of 85 item pairs with correlated errors were found when the minimum modification index was set at the critical value of 10.83. When the fit of the Icelandic model was examined, 68 pairs of items with correlated errors were discovered. In sum, the present results provide rather compelling evidence that the IBI has weak psychometric properties, although slightly improved in

the Icelandic version. Data on correlated errors from previous studies were not available to the researcher.

Although the task of improving either version of the IBI was beyond the scope of this study, preliminary steps were taken to get a rough idea of how much modification the model would need to achieve a recommendable level of fit, should future improvements be attempted. A CFI of .95 as recommended by Hu and Bentler (1999) was not possible through any means of modification, so a CFI of .90 was agreed upon. The modifications were done by running several models in succession, each time adding a pair of correlated errors from the previous model until a CFI of near .90 was obtained (see Table 15 for full results). The Icelandic model achieved a CFI of .90 after 55 correlated errors had been added to the model. When the U.S. model was modified in the same way, adding 72 pairs of correlated errors, a CFI of .90 was reached, but only barely (Table 14.). The generally recommended CFI of .95 or higher was not obtainable with the present form of the IBI, but the modified model described below (Table 15) was used for the present study's purposes.

Table 15
CFA for the IBI, Icelandic and U.S. Samples, Modified Model

Fit Measure	U.S. (N=827)	Iceland (N=720)
χ^2	2116.51	1951.37
df	1093	1110
CFI	.90	.90
RMSEA	.03	.03
SRMR	.05	.05
Largest modification index	61.81	36.41
Description of largest modification index	Cross-loading of item 23	Cross-loading of item 23

Item pairs with correlated errors are displayed in Table 16. Since modifying a model in this way model yields constantly changing modification indices, their specific value is not helpful in comparing the IBI across the two cultures. However, the item pairs identified as having correlated errors can provide valuable information regarding the conceptual organization of the IBI; if the same items and item pairs are shown as having correlated errors in the two cultures, the items can reasonably be identified as representing a latent construct common among the cultures, or at least a culturally analogous way of responding to statements regarding such a latent construct. Such information might in turn be used to improve the IBI by rewording, eliminating, or merging items. Table 16 is an abbreviated table showing the item pairs that showed up as having correlated errors in both cultures.

Table 16
*Item Pairs with Correlated Errors on U.S. and Icelandic Versions of the IBI **

22. Sometimes I can't get a fear off my mind	6. I often can't get my mind off some concern
36. It is difficult for me to do unpleasant chores	13. I usually try to avoid chores which I dislike doing
29. One should blame oneself severely for all mistakes and wrongdoings	9. Punishing oneself for all errors will prevent future mistakes
14. Those who do wrong deserve to be blamed	3. Certain people are bad or wicked and should be severely punished for their sins
30. People are disturbed not by situations but by the view they take of them	17. Nothing is upsetting in itself - only in the way you interpret it
38. Too many evil persons escape the punishment they deserve	3. Certain people are bad or wicked and should be severely punished for their sins
16. I tend to worry about possible accidents and disasters	10. I hardly ever think of such things as death or atomic war
50. I have considerable concern with what people are feeling about me	37. It is important to me that others approve of me
49. People who are miserable have usually made themselves that way	46. Man makes his own hell within himself
44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant	39. It is realistic to expect that there should be no incompatibility in marriage
37. It is important to me that others approve of me	5. I want everyone to like me
45. I can't stand to take chances.	13. I usually try to avoid chores which I dislike doing
16. I tend to worry about possible accidents and disasters	2. I worry a lot about certain things in the future

*Each item pair is presented in a row across the table
Table continues on next page

Table 16 Continued

*Item Pairs with Correlated Errors on Both U.S. and Icelandic Versions of the IBI **

29. One should blame oneself severely for all mistakes and wrongdoings	32. I get terribly upset and miserable when things are not the way I like them to be
32. I get terribly upset and miserable when things are not the way I like them to be	6. I often can't get my mind off some concern
50. I have considerable concern with what people are feeling about me	43. What others think of you is most important
32. I get terribly upset and miserable when things are not the way I like them to be	22. Sometimes I can't get a fear off my mind
49. People who are miserable have usually made themselves that way	17. Nothing is upsetting in itself - only in the way you interpret it
8. I am fairly easygoing about life	1. If I can't keep something from happening, I don't worry about it
25. I shrink from facing a crisis or difficulty	11. I avoid facing my problems
38. Too many evil persons escape the punishment they deserve	20. It is sinful to doubt the Bible
41. Immorality should be strongly punished	4. People should observe moral laws more strictly than they do
23. I hate to fail at anything	28. Frustrations upset me
36. It is difficult for me to do unpleasant chores	25. I shrink from facing a crisis or difficulty

*Each item pair is presented in a row of the table

Factor Loadings. In the U.S. model, the range of unstandardized factor loadings was -.07 to 1.64 for the whole IBI, which was also the range on the worrying subscale. For the rest of the subscales, the ranges were .15 to 1.02 (rigidity), .29 to 1.11 (problem avoidance), .34 to 1.27 (need for approval), and .88 to 1.16 (emotional irresponsibility). The range for standardized loadings across all subscales was -.03 to .76. On the subscales, the ranges were -.03 to .71 (worrying), .12 to .60 (rigidity), .22 to .59 (problem avoidance), .24 to .88 (need for approval), and .50 to .63 (emotional irresponsibility). A comparison table of the factor loadings across the two countries is provided in Table 17, grouped by subscale.

In the Icelandic model, unstandardized factor loadings, excluding the model identifier loading fixed at 1.0, ranged from .24 to 1.91 for the 50-item inventory. For the individual subscales, the unstandardized loadings ranged from .93 to 1.91 (worrying), .24

to .89 (rigidity), .47 to 1.36 (problem avoidance), .76 to 1.68 (need for approval), and .96 to 1.55 (emotional irresponsibility). When it came to the standardized loadings, the ranges were .10 to .84 for the whole IBI. For the subscales, the ranges were .28 to .64 (worrying), .10 to .52 (rigidity), .25 to .63 (problem avoidance), .41 to .84 (need for approval), and .43 to .69 (emotional irresponsibility).

Table 17
Unstandardized and Standardized Factor Loadings for the Multigroup CFA

Subscale	Item	U.S.		Iceland	
		Unstd. Loadings	Std. Loadings	Unstd. Loadings	Std. Loadings
Worrying	1. If I can't keep something from happening, I don't worry about it	1.00	0.41	1.00	0.32
	2. I worry a lot about certain things in the future	1.15	0.55	1.91	0.63
	6. I often can't get my mind off some concern	1.31	0.63	1.53	0.58
	7. I tend to become terribly upset when things are not the way I like them to be	1.64	0.71	1.81	0.62
	8. I am fairly easygoing about life	0.76	0.39	1.64	0.57
	10. I hardly ever think of such things as death or atomic war	0.45	0.18	0.93	0.28
	16. I tend to worry about possible accidents and disasters	1.12	0.49	1.60	0.50
	19. I often get excited or upset when things go wrong	1.23	0.60	1.81	0.64
	22. Sometimes I can't get a fear off my mind	1.37	0.57	1.63	0.63
	26. I feel little anxiety over unexpected danger or future events	-0.07	-0.03	1.57	0.53
	28. Frustrations upset me	0.90	0.54	1.61	0.60
	32. I get terribly upset and miserable when things are not the way I like them to be	1.53	0.67	1.64	0.60

Table continues on next page

Table 17 Continued
Unstandardized and Standardized Factor Loadings for the Multigroup CFA

Subscale	Item	U.S.		Iceland	
		Unstd. Loadings	Std. Loadings	Unstd. Loadings	Std. Loadings
Rigidity	3. Certain people are bad or wicked and should be severely punished for their sins	1.00	0.55	1.00	0.51
	4. People should observe moral laws more strictly than they do	0.75	0.53	0.36	0.28
	9. Punishing oneself for all errors will prevent future mistakes	0.65	0.43	0.67	0.43
	14. Those who do wrong deserve to be blamed	0.70	0.48	0.75	0.44
	18. A large number of people are guilty of bad sexual conduct	0.77	0.46	0.44	0.28
	20. It is sinful to doubt the Bible	0.97	0.43	0.63	0.35
	24. The fear of punishment helps people to be good	0.78	0.45	0.69	0.38
	29. One should blame oneself severely for all mistakes and wrongdoings	0.68	0.44	0.82	0.48
	33. More people should face up to the unpleasantness of life	0.40	0.26	0.40	0.26
	34. Helping others is the very basis of life	0.15	0.12	0.24	0.18
	35. There is a right way to do everything	0.82	0.43	0.56	0.30
	38. Too many evil persons escape the punishment they deserve.	1.02	0.60	0.80	0.50
	39. It is realistic to expect that there should be no incompatibility in marriage	0.46	0.29	0.18	0.10
41. Immorality should be strongly punished	0.90	0.58	0.89	0.52	
Problem avoidance	11. I avoid facing my problems	1.00	0.54	1.00	0.59
	13. I usually avoid chores which I dislike doing	0.96	0.63	1.31	0.48
	25. I shrink from facing a crisis or difficulty	0.78	0.58	1.05	0.48
	27. If something is necessary. I do it even if it is unpleasant	0.29	0.22	0.79	0.49
	31. I usually put off important decisions	1.08	0.55	1.36	0.61
	36. It is difficult for me to do unpleasant chores	1.05	0.58	1.27	0.61
	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them	1.11	0.58	1.30	0.58
	44. One should rebel against doing unpleasant things. however necessary. if doing them is unpleasant	0.44	0.27	0.45	0.30
	45. I can't stand to take chances	0.52	0.34	0.48	0.25
47. I dislike responsibility	0.70	0.44	0.96	0.50	

Table continues on next page

Table 17 Continued
Unstandardized and Standardized Factor Loadings for the Multigroup CFA

Subscale	Item	U.S.		Iceland	
		Unstd. Loadings	Std. Loadings	Unstd. Loadings	Std. Loadings
Need for approval	5. I want everyone to like me	1.00	0.68	1.00	0.53
	21. I often worry about how people approve of and accept me	1.27	0.83	1.68	0.77
	23. I hate to fail at anything	0.34	0.24	0.81	0.41
	37. It is important to me that others approve of me	1.34	0.88	1.22	0.69
	43. What others think of you is most important	0.77	0.55	1.57	0.84
	48. Although I like approval, it's not a real need for me	0.72	0.48	0.76	0.43
	50. I have considerable concern with what people are feeling about me	1.17	0.78	1.55	0.75
Emotional Irresponsibility	12. A person won't stay angry or blue long, unless he/he keeps himself/herself that way	1.00	0.56	1.00	0.45
	15. If one wants to, one can be happy under almost any circumstances	1.02	0.57	0.96	0.45
	17. Nothing is upsetting in itself - only in the way you interpret it	0.95	0.50	1.16	0.51
	30. People are disturbed not by situations but by the view they take of them	0.88	0.55	0.85	0.43
	42. There is never any reason to remain sorrowful for very long	1.03	0.55	1.19	0.49
	46. People make their own hell within themselves	0.96	0.54	1.55	0.69
	49. People who are miserable have usually made themselves that way	1.16	0.63	1.32	0.61

Residual Variances. Standardized residual variances, or error variances, in both samples were very low. The range in both samples was from .02 on several items to .07 on only one item. Item 10 had the highest error variance in both countries (Table 18).

Table 18
Standardized Residuals Produced in the Multigroup CFA, by Country

Subscale	Item	Standardized residuals	
		U.S.	Iceland
Worrying	1. If I can't keep something from happening, I don't worry about it	.05	.06
	2. I worry a lot about certain things in the future	.03	.05
	6. I often can't get my mind off some concern	.03	.04
	7. I tend to become terribly upset when things are not the way I like them to be	.04	.05
	8. I am fairly easygoing about life	.04	.04
	10. I hardly ever think of such things as death or atomic war	.07	.07
	16. I tend to worry about possible accidents and disasters	.04	.06
	19. I often get excited or upset when things go wrong	.03	.04
	22. Sometimes I can't get a fear off my mind	.04	.03
	26. I feel little anxiety over unexpected danger or future events	.05	.05
	28. Frustrations upset me	.02	.04
32. I get terribly upset and miserable when things are not the way I like them to be	.04	.04	
Rigidity	3. Certain people are bad or wicked and should be severely punished for their sins	.05	.06
	4. People should observe moral laws more strictly than they do	.03	.03
	9. Punishing oneself for all errors will prevent future mistakes	.03	.04
	14. Those who do wrong deserve to be blamed	.03	.04
	18. A large number of people are guilty of bad sexual conduct	.04	.04
	20. It is sinful to doubt the Bible	.08	.05
	24. The fear of punishment helps people to be good	.05	.05
	29. One should blame oneself severely for all mistakes and wrongdoings	.04	.04
	33. More people should face up to the unpleasantness of life	.04	.04
	34. Helping others is the very basis of life	.03	.03
	35. There is a right way to do everything	.06	.05
38. Too many evil persons escape the punishment they deserve	.04	.04	
39. It is realistic to expect that there should be no incompatibility in marriage	.04	.05	
41. Immorality should be strongly punished	.03	.04	

Table continues on next page

Table 18 Continued
Standardized Residuals Produced in the CFA, by Country

Subscale	Item	Standardized residuals	
		U.S.	Iceland
Problem avoidance	11. I avoid facing my problems	.03	.03
	13. I usually avoid chores which I dislike doing	.05	.04
	25. I shrink from facing a crisis or difficulty	.03	.03
	27. If something is necessary. I do it even if it is unpleasant	.02	.03
	31. I usually put off important decisions	.04	.05
	36. It is difficult for me to do unpleasant chores	.04	.04
	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them	.04	.05
	44. One should rebel against doing unpleasant things. however necessary. if doing them is unpleasant	.03	.03
	45. I can't stand to take chances	.03	.04
47. I dislike responsibility	.03	.04	
Need for approval	5. I want everyone to like me	.03	.04
	21. I often worry about how people approve of and accept me	.02	.03
	23. I hate to fail at anything	.04	.05
	37. It is important to me that others approve of me	.02	.03
	43. What others think of you is most important	.03	.03
	48. Although I like approval. it's not a real need for me	.04	.04
50. I have considerable concern with what people are feeling about me	.03	.03	
Emotional Irresponsibility	12. A person won't stay angry or blue long. unless he/he keeps himself/herself that way	.04	.04
	15. If one wants to. one can be happy under almost any circumstances	.04	.04
	17. Nothing is upsetting in itself - only in the way you interpret it	.05	.05
	30. People are disturbed not by situations but by the view they take of them	.03	.04
	42. There is never any reason to remain sorrowful for very long	.04	.05
	46. People make their own hell within themselves	.04	.04
	49. People who are miserable have usually made themselves that way	.04	.04

Factor Correlations. Despite the poor fit with the five-factor solution, correlations among factors in the U.S. model were somewhat consistent with those obtained by Bridges and Sanderman (2002), and in most cases, the difference in correlation was less

than .1 (see Table 19). The biggest departure from Bridges and Sanderman's (2002) study was the correlation of the worrying factor with problem avoidance, which was .41 in the present study but .18 in Bridges and Sanderman's (2002) study. The correlations between the factors for the U.S. and Icelandic samples are shown and compared to Bridges and Sanderman's (2002) translation, and Koopmans et al.'s (1994) original Dutch study in Table 19.

Table 19
Current Correlations of Subscales Compared with Bridges & Sanderman (2002) and Koopmans et al. (1994)

Scale	Worrying		Rigidity		Need for Approval		Problem Avoidance	
	U.S./ Iceland/ U.S.'02 / Dutch '92	U.S.'02 /	U.S./ Iceland/ U.S.'02 / Dutch '94	U.S.'02 /	U.S./ Iceland/ U.S.'02 / Dutch '94	U.S.'02 /	U.S.'02 / Dutch '94	
Worrying	–		–		–		–	
Rigidity	.32 / .38 / .29 / .16		–		–		–	
Need for Approval	.52 / .58 / .46 / .45		.16 / .31 / .29 / .14		–		–	
Problem Avoidance	.41 / .39 / .18 / .25		.27 / .17 / .18 / .09		.41 / .41 / .23 / .19		–	
Emotional Irresponsibility	.20 / .23 / .19 / .15		-.23 / -.21 / .19 / -.06		.03 / .06 / -.01 / -.03		.07 / .08 / -.03 / -.07	

Correlations between the factors were quite similar to the ones obtained for the U.S. version, and the correlation between the need for approval and problem avoidance subscales was exactly the same. The biggest difference between the countries was a .15 difference between the rigidity scale and need for approval, and a .1 difference between the rigidity and problem avoidance scales, the Icelandic correlations being lower in both cases. Overall, the correlations between subscales followed a similar pattern in both the U.S. and Iceland. The biggest correlation between subscales in both countries was between the worrying scale and need for approval, followed by moderate correlations between the problem avoidance subscale and two other subscales, need for approval and worrying (Table 19).

Section Summary. Two confirmatory factor analyses were carried out to examine the fit of the U.S. and Icelandic versions, respectively, to the U.S. model of the IBI presented by Bridges and Sanderman (2002). The chi-square and CFI conflicted with the RMSEA and SRMR, with the former two showing poor fit and the latter showing acceptable fit. Overall psychometric properties of the IBI were found to be weak, the Icelandic model performing marginally better than the U.S. model. Modified models that included pairs of correlated errors in both samples produced a better-fitting model, with improvements only reaching a CFI of .90 in both countries. Correlations among subscales in both Iceland and the U.S. were moderate overall, and consistent with each other.

Invariance Testing of the IBI

One of the purposes of the current study was to determine whether or not the Irrational Beliefs Inventory measures the construct of irrational beliefs in an equivalent way across two countries, Iceland and the United States. To test whether such equivalence existed, multigroup CFA was used, implemented using Mplus statistical software, version 4 (Muthén & Muthén, 2008). Multigroup CFA approach allows a detailed comparison of all structural parameters of a measurement model across groups, provided that the factor structure of the model is a good enough match with the relevant theory to make inferences possible (Brown, 2006). The process of comparing the equivalence of the construct in this way is called an invariance test, invariance referring to sameness or equivalence. In an invariance test, a change in chi-square relative to change in degrees of freedom for nested models (e.g., Model 1 with loadings free to vary across countries vs. Model 2 with loadings constrained to be equal across countries) indicates that items may be performing differently across the respective groups, which in

turn can give clues to which items need modification. In turn, a non-significant difference in chi-square indicates invariance or equivalence, meaning that the item in question appears to perform comparably in both samples. Differences, or non-invariance, between samples can mean a number of things in the current study: The items on the test might have different meaning or be related to different concepts in the different cultures, a bias might exist in the translation of an item, and so on. If statistical invariance is established for the IBI when comparing the Icelandic and U.S. samples, it is tenable that the new Icelandic version of the IBI is equivalent to the U.S. version.

In the current study, maximum likelihood estimation was used. The covariance matrix of the 50 variables (items) was used for input. Due to the poor fit and overall weak psychometric properties of the original model, the modified model with 72 correlated errors on the U.S. side and 55 on the Icelandic side was used for the baseline model of the invariance test. The initial step was to test a baseline model between the Icelandic and U.S. samples. A baseline model contains no statistical constraints on any of the model's parameters; thus, a baseline or starting point was provided to allow testing for invariance of various parameters. After the baseline model, equality of factor loadings was tested in an omnibus test, followed by individual tests of loadings to determine where the differences were located. Following tests of the factor loadings, invariance tests for the intercepts, residuals, factor variances, and factor covariances were conducted (See Table 20).

Upon examination of the baseline model, non-equivalence of the factor loadings between the two countries was discovered, $\Delta\chi^2=443.72$, $\Delta df = 45$, $p < .001$. To locate the specific source of the difference, an examination of the factor loadings was taken as

the next step. Since there are 50 items on the IBI, with five factors for which fit was tested, the factor loading equality constraint resulted in a change in 45 degrees of freedom (i.e., one item within each factor was constrained to 1.0 for identification purposes). The breakdown of the models is listed in Table 20.

Table 20
CFA Models Used in the Cross-Cultural Invariance Testing of the IBI

Model	χ^2	df	$\Delta\chi^2$	Δ df	CFI	RMSEA	SRMR
1. Baseline model with no equality constraints	4067.87	2203	-	-	.90	.03	.05
2. Loadings (45) constrained to be equal	4511.51	2248	443.72	45	.88	.04	.06
2a. Loadings (24) constrained to be equal	4114.59	2227	46.72	24	.90	.03	.05
3. Intercepts (24) constrained to be equal	6485.14	2251	2370.55	24	.77	.05	.07
3a. Intercepts (11) constrained to be equal	4177.71	2238	63.12	11	.90	.03	.05
4. Residuals constrained to be equal	4531.82	2288	354.11	50	.88	.04	.06
5. Factor variances constrained to be equal	4586.65	2293	54.83	5	.88	.04	.06
6. Factor covariances constrained to be equal	4619.17	2303	32.52	15	.88	.04	.06

When the loadings were examined one by one, 21 items were found with $\Delta\chi^2$ values significantly exceeding the critical value of 10.83 at $p < .001$, and thus there was evidence that these 21 items were functioning differently across the countries. The items are shown in Table 21, arranged in descending order by the change in chi-square.

Table 21

Items Contributing to Difference Between the Icelandic and U.S. Samples When Factor Loadings Were Constrained to be Equal

Subscale	Item description	Unstd. loadings			$\Delta\chi^2$
		U.S.	Iceland	Diff. in loading	
Worrying	26. I feel little anxiety over unexpected danger or future events	-0.07	1.41	1.48	79.72
Need for approval	43. What others think of you is most important	0.77	1.75	0.98	47.72
Need for approval	23. I hate to fail at anything	0.36	1.06	0.70	33.40
Need for approval	21. I often worry about how people approve of and accept me	1.29	2.27	0.98	30.31
Problem avoidance	27. If something is necessary, I do it even if it is unpleasant	0.34	0.86	0.54	25.47
Need for approval	50. I have considerable concern with what people are feeling about me	1.28	2.05	0.77	20.14
Worrying	8. I am fairly easygoing about life	0.74	1.52	0.78	18.30
Emotional irresponsibility	46. People make their own hell within themselves	.95	1.55	0.59	11.77
Rigidity	34. Helping others is the very basis of life	0.14	0.46	.32	9.77
Worrying	2. I worry a lot about certain things in the future	1.10	1.78	.68	8.62
Emotional irresponsibility	17. Nothing is upsetting in itself - only in the way you interpret it	0.74	1.17	.43	8.26
Rigidity	39. It is realistic to expect that there should be no incompatibility in marriage	0.53	0.17	.38	7.67
Rigidity	18. A large number of people are guilty of bad sexual conduct.	0.92	0.52	.40	6.84
Problem avoidance	31. I usually put off important decisions	1.05	1.44	.39	6.12
Worrying	10. I hardly ever think of such things as death or atomic war	0.33	0.76	.43	6.09
Rigidity	4. People should observe moral laws more strictly than they do	0.90	0.57	.33	5.28
Rigidity	20. It is sinful to doubt the Bible	1.35	0.87	.48	5.06
Problem avoidance	47. I dislike responsibility	0.67	0.98	.31	5.04
Problem avoidance	45. I can't stand to take chances	0.71	0.49	.22	3.37
Emotional irresponsibility	42. There is never any reason to remain sorrowful for very long	0.93	1.20	.27	3.23
Rigidity	38. Too many evil persons escape the punishment they deserve	1.08	0.84	.24	2.57

Four out of eight items for the need for approval subscale functioned significantly different between Iceland and the U.S. This finding corresponds with qualitative data from Phase 1, during which items from the need for approval scale were found particularly difficult to translate. The remaining 24 items on which the factor loadings invariance test did not display a difference are displayed in Table 22.

Table 22
Items with Equal Loadings across the Samples, Arranged by Change in Chi-square

Subscale	Item description	U.S.	Iceland	Diff. in loading	$\Delta\chi^2$
Emotional irresponsibility	49. People who are miserable have usually made themselves that way	1.05	1.31	.26	2.56
Worrying	32. I get terribly upset and miserable when things are not the way I like them to be	1.54	1.15	.39	2.51
Worrying	16. I tend to worry about possible accidents and disasters	1.06	1.40	.34	2.31
Need for approval	48. Although I like approval, it's not a real need for me	.78	.98	.20	2.31
Worrying	28. Frustrations upset me	.89	1.16	.26	2.08
Problem avoidance	13. I usually try to avoid chores which I dislike doing	.99	1.18	-.19	1.47
Problem avoidance	25. I shrink from facing a crisis or difficulty	.90	1.04	-.13	1.17
Rigidity	33. More people should face up to the unpleasantness of life	.43	.55	-.13	.97
Emotional irresponsibility	30. People are disturbed not by situations but by the view they take of them	.74	.85	-.12	.81
Worrying	7. I tend to become terribly upset when things are not the way I would like them to be	1.64	1.41	.23	.71
Worrying	6. I often can't get my mind off some concern	1.26	1.41	-.15	.46
Worrying	22. Sometimes I can't get a fear off my mind	1.37	1.54	-.17	.44
Rigidity	29. One should blame oneself severely for all mistakes and wrongdoings	.73	.83	-.10	.44
Problem avoidance	44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant	.44	.37	.07	.41
Rigidity	41. Immorality should be strongly punished	1.08	1.17	-.10	.28
Worrying	19. I often get excited or upset when things go wrong	1.27	1.39	-.11	.25
Rigidity	14. Those who do wrong deserve to be blamed	.73	.79	-.06	.19
Rigidity	9. Punishing oneself for all errors will prevent future mistakes	.66	.71	-.05	.12

Table continues on next page

Table 22 Continued
Items with Equal Loadings across the Samples, Arranged by Change in Chi-square

Subscale	Item description	U.S.	Iceland	Diff. in loading	$\Delta\chi^2$
Rigidity	35. There is a right way to do everything	.92	.73	.19	.12
Need for approval	37. It is important to me that others approve of me	1.32	1.28	.05	.11
Problem avoidance	36. It is difficult for me to do unpleasant chores	1.07	1.11	-.05	.09
Rigidity	24. The fear of punishment helps people to be good	.78	.83	-.04	.08
Problem avoidance	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them	1.23	1.25	-.02	.02
Emotional irresponsibility	15. If one wants to, one can be happy under almost any circumstances	.98	.97	.01	.00

After the invariance test of factor loadings had determined that 21 items functioned differently between the two countries, the next level of analysis was an examination of the remaining 24 items' intercepts, minus the five items used as model identifiers. Intercepts reflect group means and show how the different groups endorse the items. When the intercepts of the 24 items were analyzed, a total of 13 items were found to have significantly different group means across the countries. Items included in the intercept model are shown in Table 23, and intercepts for items used in previous tests in which the loadings were found to be significantly different between countries are displayed in Table 24.

Table 23
Items Used in the Intercept Invariance Test, Arranged by Change in Chi-square

Subscale	Item	Intercept value		Intercept difference	$\Delta\chi^2$
		U.S.	Iceland		
Rigidity	14. Those who do wrong deserve to be blamed	3.49	2.12	1.37	^a
Rigidity	24. The fear of punishment helps people to be good	3.23	2.33	0.90	206.83*
Need for approval	48. Although I like approval. it's not a real need for me	2.66	1.97	0.69	150.43*
Problem avoidance	44. One should rebel against doing unpleasant things. however necessary. if doing them is unpleasant	2.22	1.74	0.48	121.01*
Problem avoidance	13. I usually avoid chores which I dislike doing	3.17	2.48	0.69	101.05*
Worrying	22. Sometimes I can't get a fear off my mind	3.00	4.04	1.04	95.31*
Rigidity	9. Punishing oneself for all errors will prevent future mistakes	2.13	1.67	0.46	69.68*
Worrying	7. I tend to become terribly upset when things are not the way I like them to be	3.03	3.75	0.72	54.41*
Emotional Irresponsibility	15. If one wants to. one can be happy under almost any circumstances	2.29	1.93	0.36	36.35*
Emotional Irresponsibility	49. People who are miserable have usually made themselves that way	2.72	3.16	0.44	34.03*
Worrying	6. I often can't get my mind off some concern	3.63	4.03	0.40	24.55*
Worrying	28. Frustrations upset me	3.79	3.45	0.34	21.99*
Rigidity	33. More people should face up to the unpleasantness of life	3.13	3.41	0.28	20.49*
Rigidity	29. One should blame oneself severely for all mistakes and wrongdoings	2.18	1.98	0.20	10.4
Worrying	16. I tend to worry about possible accidents and disasters	3.00	2.74	0.25	9.01
Emotional Irresponsibility	30. People are disturbed not by situations but by the view they take of them	2.65	2.53	0.12	4.86
Problem avoidance	36. It is difficult for me to do unpleasant chores	2.78	2.66	0.12	3.98
Rigidity	35. There is a right way to do everything	2.80	2.65	0.15	3.93
Problem avoidance	40. I often spend more time trying to think of ways of getting out of things than it would take me to do them	2.44	2.55	0.11	2.64
Need for approval	37. It is important to me that others approve of me	3.12	3.04	0.08	1.83
Rigidity	41. Immorality should be strongly punished	2.77	2.69	0.08	1.20
Worrying	32. I get terribly upset and miserable when things are not the way I like them to be	2.81	2.88	0.07	0.79
Worrying	19. I often get excited or upset when things go wrong	3.32	3.39	0.07	0.69
Problem avoidance	25. I shrink from facing a crisis or difficulty	2.49	2.46	0.03	0.31

*Statistically significant with critical value at 10.83 at 1 df

^aNo convergence, too many iterations needed to process item successfully, difference in intercepts used to assume significant difference.

Table 24
Intercepts for the 21 Items Not Used in the Intercept Invariance Test Due to Unequal Factor Loadings

Subscale	Item	Intercept value		Intercept difference
		U.S.	Iceland	
Rigidity	20. It is sinful to doubt the Bible	2.67	1.37	1.30
Need for approval	50. I have considerable concern with what people are feeling about me	2.90	2.01	0.89
Need for approval	21. I often worry about how people approve of and accept me	3.13	2.34	0.79
Need for approval	23. I hate to fail at anything	3.91	3.13	0.78
Worrying	8. I am fairly easygoing about life	2.16	2.79	0.63
Need for approval	43. What others think of you is most important	2.21	2.77	0.56
Worrying	2. I worry a lot about certain things in the future	3.79	3.30	0.49
Problem avoidance	45. I can't stand to take chances	2.12	2.56	0.44
Rigidity	18. A large number of people are guilty of bad sexual conduct	3.13	3.50	0.37
Rigidity	39. It is realistic to expect that there should be no incompatibility in marriage	2.26	1.93	0.33
Rigidity	38. Too many evil persons escape the punishment they deserve	3.29	3.61	0.32
Problem avoidance	47. I dislike responsibility	1.96	2.26	0.30
Emotional Irresponsibility	17. Nothing is upsetting in itself - only in the way you interpret it	3.02	2.74	0.28
Problem avoidance	31. I usually put off important decisions	2.51	2.72	0.21
Emotional Irresponsibility	46. People make their own hell within themselves	2.72	2.52	0.20
Worrying	26. I feel little anxiety over unexpected danger or future events	2.92	3.10	0.18
Problem avoidance	27. If something is necessary. I do it even if it is unpleasant	2.09	2.27	0.18
Rigidity	34. Helping others is the very basis of life	3.98	3.83	0.15
Emotional Irresponsibility	42. There is never any reason to remain sorrowful for very long	3.02	2.97	0.05
Worrying	10. I hardly ever think of such things as death or atomic war	3.00	2.96	0.04
Rigidity	4. People should observe moral laws more strictly than they do	3.70	3.67	0.03

To test the equality of reliabilities across groups, an invariance test of residuals was conducted. Admittedly, such a test is inherently weak as a stand-alone test due to the low degree of equivalence between the factors, but complemented the invariance test for the purposes of the current study. A similar caveat follows the invariance test of factor

variances and factor covariances, since they depend on the factor loadings to be equal for the results to be of substantive use (Brown, 2006). The chi-square test revealed significant differences in the model for both the factor variances and covariances, but due to the weak fit and high degree of correlated error in the original model, and the high degree of modification needed to reach acceptable fit indices, it is extremely difficult to use the factor variances and covariances to reach definitive conclusions about the equivalence of the factors and their covariances across Iceland and the U.S. Hence, it could be argued that an in-depth analysis of the variances and covariances would be an example of an artificially fitted, or *overfitted* model, and thus not the soundest of procedures when it comes to discussing factorial invariance (Brown, 2006).

In summary, the psychometric analyses of the IBI tentatively supports partial measurement invariance for the 11 items that demonstrated non-significant differences on their loadings and intercepts. These results along with the results of the translation, adaptation, and cognitive interview phases of the study will be discussed in the next chapter.

Chapter Five

Discussion

The purposes of the study were to translate and adapt the Irrational Beliefs Inventory from U.S. English to Icelandic, and then compare the psychometric properties of the U.S. and Icelandic versions. These purposes were accomplished in three major phases, a translation and adaptation phase, a qualitative cognitive interview and cultural validation phase, and a psychometric evaluation phase. This chapter contains six sections. The first section revisits the construct of irrational beliefs as conceptualized in the theoretical framework of Rational Emotive Behavior Therapy. In the second section, major findings are summarized. Findings from the translation and adaptation phases are discussed in detail in the third section, and the fourth section features discussion on the psychometric evaluation phase. In the fifth section, results from all three phases are synthesized, and the significance of the current study within the scope of cross-cultural research is analyzed. In the sixth and final section, suggestions for future research are offered.

Theoretical Background of the IBI and the Current Findings

When an instrument measuring a psychopathological construct becomes a subject of cross-cultural validation, findings can be interpreted from several perspectives. One perspective is at the individual level, in which the cross-cultural comparability of the measured construct is examined by analyzing individuals' scores on the questionnaire. Another perspective is a broader socio-cultural one, which invites reflection on factors

such as political systems, literacy and poverty levels, demographics, and access to healthcare and information, that may play into an individual developmental history. Yet another perspective is methodological, where the instrument itself and methods to translate and adapt it are critically evaluated. The original main focus of the present study was on the translation, adaptation, and factorial validation of the IBI, that is, how well the Icelandic five-factor model of the IBI matched the five-factor model of its U.S. counterpart. However, due to inconclusive results from the psychometric data, a social-contextual approach will be used to complement the findings.

The Irrational Beliefs Inventory (IBI) is rooted in the theoretical framework of Rational Behavior Emotive Therapy (REBT), which evolved when Albert Ellis combined insights from psychodynamic theory with Learning Theory (Ellis, 1962). The main idea behind REBT is that emotions, which help the human species survive threats and reproduce, can be interpreted in an illogical or inappropriate context and thereby create *irrational* beliefs about the world and him/herself (Ellis, 1962). Ellis maintained that most psychological disturbances are caused by a dogmatic and absolutistic thinking style that mediates stressful, exaggerated, and or/prolonged emotional responses to situations where such responses are not helpful (Ellis, 1962). Specifically, the human animal's ability for symbolic communication (e.g., language) makes it able to create memories of intrinsically stressful or pleasing events, accompanied with the corresponding emotions, when the original emotion-eliciting event has long passed (Ellis, 1962). This is in contrast to what is observed in many other places of the animal kingdom. For example, when antelopes run for their lives from a lion, and the lion catches one, the antelopes can be seen calmly resuming grazing with no apparent sign of lingering stress, only minutes

after a lion started eating one of them. Had this happened to a human group, it is not unreasonable to expect long-term repercussions with the survivors, such as recurrent nightmares, feelings of guilt, and prolonged grief, perhaps sustained by elaborate remembrance rituals during which emotionally conditioned symbols (e.g., photographs or artifacts that had belonged to the dead) are treated with almost the same care as if they were in fact the deceased person. This ability to recreate and relive events through language and memory also gives one an ability to put past experiences in a present context that is no longer applicable. This context is provided by the person's individual developmental history, as well as the language and customs of the culture (Ellis, 1962). The positive side of this ability may perhaps be seen in some of humanity's most marvelous achievements, such as romantic poetry and music, which can take a person to wonderful faraway places by its conditioned evocations alone (Morris, 1969). The downside of this ability is that illogical interpretations of past events can become emotionally conditioned to the memory of the event and thus create faulty generalizations that gradually become irrational beliefs and thus trap people in mistakes of their past (Dryden, DiGiuseppe, & Neenan, 2003; Ellis, 1962; Ellis & Harper, 1997).

A person's culture and language undeniably contribute to the development of beliefs, by providing the framework for the developing individual's experiences and social interactions (Bronfenbrenner, 1979, 1994; Morris, 1969). This in turn means that a person is dependent on the culture in order to form healthy beliefs, and the survival of the culture in turn depends on the opportunities it provides to its people to learn adaptive interpretations to stressful situations. Thus, REBT should, in theory, be relevant for the important issues facing an increasingly diversifying and interconnected world. In order to

provide fair measurement methods of irrational beliefs and their development, the theories we use to describe such beliefs must be applicable universally to all members of our species; however, our measurement methods must be applicable locally, to members of that particular culture.

Overview of Findings

The findings of the current study were produced in the course of three phases of the research: a translation and adaptation phase, a qualitative interview phase, and a psychometric phase, all of which will be discussed in greater detail in their respective subsections below. The translation and adaptation phase was accomplished in three steps. First, the IBI was translated by three translators, and the translation polished via email communication until all disagreements had been settled. The next step was to translate the IBI back into English (a process known as back-translation) by two back-translators, and their versions compared to each other and the forward-translated version. The final step of phase one was a review by a committee of three, who evaluated how well the IBI had been adapted to the Icelandic culture.

Phase two was a qualitative cognitive interview phase, in which small samples of participants in both the U.S. and Iceland responded to the IBI and then were asked questions about the meaning of individual items. Translations and interviews alike provided hints that items describing values (e.g., ethics, religion), as well as the general tenor of the whole instrument would prove a challenge to adapt to the Icelandic culture.

Phase three was a large-scale psychometric evaluation of the IBI, in which a Confirmatory Factor Analysis (CFA) was used to analyze the factorial validity of the IBI and cross-culturally evaluate responses from a 827-person U.S. sample and a 720-person

Icelandic sample. The psychometric data obtained in the current study did not support the existing five-factor model of the U.S. version of the Irrational Beliefs Inventory (Bridges & Sanderman, 2002), and the assumption of equality between the two versions of the IBI was not supported. Thus, the findings of the current study indicate the IBI in its current form is not a suitable research tool, and that a revision or a complete reassembly is in order. In a study published after data collection in the current study was underway, Terjesen et al. (2009) found that most psychometric instruments measuring irrational beliefs have incomplete or conflicting psychometric information (e.g., lack of manuals, lack of consensus regarding number and nature of factors). Another factor to consider in interpreting the current psychometric results for the IBI is the emerging changes in REBT theory itself. If REBT theory is indeed in a state of development, the validity of any study attempting to assess irrational beliefs is potentially jeopardized. Even with a theory in a stable developmental phase, the measurement models of the constructs in each country would need to have acceptable and comparable fit. If the model does not fit well psychometrically, its operationally defined link to the corresponding theory is weak. In the current study, CFA models of the Irrational Beliefs Inventory for Iceland and the U.S. were compared, and the overall fit of both models was weak, which limits cross-cultural comparisons of the constructs underlying the IBI. The findings do, however, shed some light on the process of cross-cultural adaptation of psychometric tests, in particular of the kind that measure broad personality-related constructs, such as irrational beliefs. Threats to the construct validity of the IBI and its psychometric siblings (see e.g., Terjesen, 2009) will be discussed in greater detail in the sections below, as well as in sections specifically discussing significance and limitations of the current study.

Phase 1: Translation, Back-Translation, and Adaptation of the IBI

Most guidelines on translation and adaptation place an emphasis on four steps of the process, forward-translation, back-translation, cultural adaptation, and psychometric analysis (e.g., Hambleton, 2005; Sireci, Yang, Harter, & Ehrlich, 2006). During this study, the first three general steps of the process overlapped considerably, because the forward-translators and back-translators all provided valuable comments regarding cultural relevance. Thus, translation and adaptation insights were found useful as qualitative data. This may have been helped by the fact that everyone involved in the forward-translations and back-translations had training in psychology or a related field, and had all lived and studied in the United States. In fact, the forward-translators and back-translators provided a far higher number of comments on cultural relevance than the evaluation committee did. This finding is consistent with recommendations that the translators be proficient in the original language (Hambleton, 2005), but it also underscores the benefits of having forward-translators and back-translators with firsthand knowledge of the academic discipline involved, as well as the culture in which the language is used (van de Vijver & Poortinga, 2005). This benefit is more subtle than just knowledge of the language, because firsthand knowledge of a culture gives subtle yet important insights into cultural context, which in turn reveals valuable sources of meaning that go untapped by those who have never lived in the culture. This includes, but is not limited to, etymology of idioms and their shades of meaning, appropriate tone of conversation, and local views on social matters such as religion and crime and punishment. IBI item 48: "Although I like approval, it's not a real need for me" is an example of items that needed both translation skills and cultural-contextual skills, due to

difficulties in translating the word "approval". In the U.S., approval can mean an agreement, such as approval of a budget, but can also be a general value judgment of a person. In Icelandic, approval refers largely to agreement, such as approval of a budget or plan, but when translated literally from the IBI and used as a value judgment of a person, it sounded strange to all translators. Instead of "approval", Icelandic equivalents of terms such as "liking", "caring for" and "respect" were chosen, because they all represent the U.S. idea of approval, and also conform to the basic REBT tenet that a rigid need to be liked by everyone contributes to stress due to its inherent futility. The researcher and translators all agreed that the concept "approval of" a whole person is quite prevalent in the U.S., but is not used in the same way in Icelandic. Apart from this study, no empirical findings on the cultural differences of the term approval are known to exist, but the decision was made with high agreement between the researcher and translation team. One of the forward-translators was a linguist with a Ph.D. from a U.S. university, and he was consulted on this and other matters related to Icelandic-specific linguistic issues. In future cases of contextual ambiguity, advice of linguists with shared cultural backgrounds of the cultures in question is recommended, if available.

The cultural review committee did not yield nearly as many comments about the whole translation and adaptation process as was expected beforehand, perhaps due to the extensive review given during translation and back-translation, but the cultural review process is still deemed essential to the translation and adaptation work. Regardless of the work done by translators and adaptors, the cultural reviewers provided a fresh pair of eyes who had not seen the original survey. Obviously, the makeup of translation teams and review teams will vary from study to study, and the more review a survey undergoes,

the better. Ideally though, a cultural review committee should consist of experts in the field (Hambleton, 2005; Sireci, Yang, Harter, & Ehrlich, 2006), but a reviewer who is neither familiar with the field in question nor the original language might be a good addition to an adaptation team, to see whether the survey's language makes sense to a layperson who belongs to the group for which the survey was intended.

Phase 2: Cognitive Interviews and Cultural Validation of the IBI

An informal discussion group session was conducted among 21 undergraduates in the U.S. The session was different from a focus group session in the sense that the researcher did not play an active role as a moderator, but simply asked students to respond to the IBI as they would with any other questionnaire, but make notes of any items that sounded strange for any reason at all. Most comments from the group related to wording on items, and potentially outdated meaning. An example was for item 3: "Certain people are bad or wicked and should be severely punished for their sins." Whereas the word "wicked" in the context of the IBI refers to a malignant nature, some members of the discussion group maintained that the word "wicked" had become a positive exclamation, sometimes uttered instead of words like "cool" or "awesome". Others indicated that "wicked" was archaic even though they understood what was meant. The word "severely" was also deemed by a couple of members of the group to be archaic (e.g., in the context of "severely punished"). As another example, in item 10: "I hardly ever think of things such as death or atomic war", the term "atomic" was found archaic by the group. Instead, some students suggested "nuclear", and other suggested that a threat of nuclear war was not present at all anymore, and suggested terrorism as a more likely world threat. In the case of an instrument such as the IBI, it is recommended that topical

threats be abandoned altogether, and a timeless fear (e.g., death, accident, disease, etc.) be used instead. A larger lesson from such unstructured discussion groups can be that they need to be conducted regularly with any given instrument, to get a feel for how colloquial terms change over time, and to observe the participants' conversation for insight into subtle shades of meaning. A potential limitation to the group discussion session was that fewer and fewer students commented on the IBI as the session wore on. When a survey is as long as the IBI (i.e., 50 items), this might be expected. A possible solution is to conduct more discussion groups with each responding to fewer items. A separate discussion group might also be used to gauge for a more gestalt look-and-feel approach.

Cognitive interviews yielded largely similar results as the group session, only more extensive. Since the cognitive interviews were conducted one-on-one and in a more relaxed environment, it was possible to ask more about each item and ask more in-depth questions. All in all, the advantages of the group session in this study were that the researcher could watch the group discuss the test items in their own language and on their own terms without any direction or manipulation. The cognitive interviews were, however, more helpful in obtaining more in-depth feedback about the items.

The discussion group and cognitive interviews provided a wealth of information about potential shortcomings of the IBI. This raised a dilemma for the researcher. On one hand, it could have been possible to modify the U.S. version of the IBI so as to make it the best instrument possible for adaptation; on the other hand, modifications to the U.S. version would have reduced comparability with existing data provided by Bridges and Sanderman (2002) with no guarantee of improved psychometric properties. The solution

was to take existing data of the IBI at face value and make no material changes to the U.S. version at this time.

Due to time constraints, a group discussion was not possible in Iceland, but cognitive interviews were conducted, which provided insights that were used to edit the IBI. Participants in the U.S. mostly commented on wording, but the Icelandic participants tended to comment more on meaning. Comments fell into three main content categories as analyzed by the researcher: specificity, meaning, and personal values. Often, items needed to be worded more specifically on the Icelandic IBI, as Icelandic participants appeared to be less inclined than their U.S. counterparts to think of items generally, despite having instructions to that effect. For example, on item 15: "If one wants to, one can be happy under almost any circumstances", some Icelandic participants commented that certain circumstances are inherently unpleasant, citing death of a loved one or staying in a hospital as examples. In this context, the term "circumstances" appeared to have a more general meaning among U.S. participants, and when translated to Icelandic, some interviewees remarked that they understood the term for "circumstances" as a reference to specific events, such as a car accident or death of a loved one. Thus, without a specific example of the "circumstance" or "situation" given in the item, item 15 became vague to the Icelanders. Another example of how cognitive interviews in Iceland provided data for edits is in item 2, "I worry a lot about certain things in the future", to which the operative term "often" was added ("I often worry a lot about certain things in the future..."), because the original wording made the item sound too focused on the present moment and did not come across as describing a general tendency.

The strong tendency of the Icelandic cognitive interviewees to view the items on the IBI as situation-specific provides an opportunity for further research and consideration when adapting or assembling a survey intended for Iceland. Spielberger, Moscoso, and Brunner (2005) argued that people in collectivist cultures seem to explain their own behavior more in terms of situations, whereas people in individualist cultures tend to use trait-specific explanations more often. Thus, in light of participants' common calls for more specificity during cognitive interviews and administration in Iceland, it is reasonable to expect that psychological tests intended for Icelanders, whose culture is arguably more collectivist than the U.S., might need to feature more situation-specific items than their U.S. counterparts.

Another example of comments related to meaning manifested itself on item 13: "I usually try to avoid chores that I dislike doing". In Icelandic, the word "chore" translates to "work I need to do", or "work of duty" more literally, which loses meaning in Icelandic, because the operative term "need" in the Icelandic language renders one's opinion of the chore irrelevant. Thus, the item appeared to confuse some participants. The researcher's solution was to reword the item to emphasize procrastination as opposed to completely avoiding the task. Yet another possibility could have been to rewrite the item completely to read as a belief rather than as a reaction (e.g., "I think that boring chores/tasks) are best put back or completely avoided if possible".

A third example of differences in interpretations between the countries was for items that pertained to core values (e.g., religion, views on crime and punishment). This topic area was considerably subtler and more difficult to translate into meaningful statements in Icelandic. All the participants in the Icelandic cognitive interviews, as well

as forward-translators, back-translators, and cultural reviewers, remarked that the IBI had a "religious undercurrent" or "religious overtone". In other words, the Icelandic participants found the IBI to have strong religious, and sometimes even fundamentalistic, overtones. Some participants in the actual study on the Iceland side commented that the IBI had an "American tone" or "religious feel", as the translation and adaptation teams had, despite efforts to tone down the perceived religious feel of the IBI after the translation and cultural review phase. A couple of participants even asked the researcher whether the IBI was intended to measure religiosity. In contrast, no participant in the U.S. raised objections to the religious items on the IBI. A possible explanation for this notable difference may lie in the different role of religion in the U.S. and Iceland. According to aggregated data from the World Values Survey collected in 1982, 1990, and 1999, 72.2% of Icelanders described themselves as "religious", whereas 83.5% of Americans did (World Values Survey, 2005). In contrast, 27.8% of Icelanders described themselves as either "not religious" or "a convinced atheist" whereas 16.4% of Americans did. However, claiming to be spiritual or religious is not the same as actively pursuing a religious lifestyle. In Iceland, just under 2% of the population reported going to church once per week, and less than 1% reported going more than twice per week. In the U.S., a little over 29% reported going once per week and 14% reported going more than once per week (World Values Survey, 2005). The simplest way to reconcile the cognitive interview findings and the World Values Survey data may be that, while many Icelanders claim to have a spiritual side to them, it may simply not manifest itself in the culture and lifestyle (e.g., attending services, or reading scriptures literally). In turn, this would mean that the researcher's decision to tone down the religious overtone of the IBI was

warranted, as religion does not appear to be a core value in Iceland. The words are the same when translated, but their emotional weight is not, so it may be worth a try to attempt to identify a core value in Iceland that corresponds to religion in the U.S. Once that has been found, items focusing on religion can either be rewritten or replaced altogether with the new items that measure Iceland-specific core values. Then, these items need to be worded in a way that discriminates between rational and irrational manifestations of the corresponding values on a five-point Likert scale. Perhaps the best idea would be to simply write a new inventory from scratch, based on local Icelandic surveying. This is a potential avenue for future research.

In sum, cognitive interviews in both countries revealed both subtle and profound differences in meaning in items of the IBI, and thus foreshadowed considerable problems with the IBI which were encountered during the quantitative analysis and will be discussed in the following section. In the U.S., most information gained from cognitive interviewing related to wording, whereas comments on the Iceland side were mostly about meaning and tenor. Comments of the translation and adaptation team somewhat mirrored those of cognitive interviewees, although less editing was needed as the process of adaptation wore on. A limitation is that time constraints prevented extensive piloting of the Icelandic IBI after the cognitive interviews had been conducted, but making more changes to the Icelandic IBI might have been risky, as some items' meaning had already been modified considerably from their U.S. versions. A suggestion for future adaptations is to share feedback from participants with the forward- and back-translators, as well as the culture review group, and then pilot test again.

Phase 3: Psychometric Analysis of the IBI

Based on the results of an initial confirmatory factor analysis (CFA), the IBI's overall psychometric performance was weak in both the U.S. and Iceland, with a CFI of only .76 in both countries. Such a weak performance raises serious questions about the psychometric properties of the IBI, so the models were modified by adding parameters estimating the correlations between pairs of errors for a number of item pairs. A total of 85 item pairs with correlated errors were added to the U.S. model and 68 pairs were added to the Icelandic model. This modification resulted in a CFI of .90 in both countries, which is considered marginally acceptable by some (e.g., Brown, 2006), but still below the .95 recommended by others (Hu & Bentler, 1999). The relative fit indices improved slightly; the RMSEA in the modified model was .03 for both countries, and the SRMR was .05, both of which are considered acceptable fit (Hu & Bentler, 1999). Absolute levels of interpreting fit indices are not necessarily recommended, due to the nature of constructs and their often-debated relationship to inferential statistics (Brown, 2006; Rigdon, 1996), but when the fit indices conflict, the most reasonable conclusion is that the fit is not good (Brown, 2006).

While factorial validity was low, the IBI was relatively free of items cross-loading on a subscale other than specified by the model. Only item 23 ("I hate to fail at anything") had a loading on a second factor. It was originally on the need for approval subscale, but had a second loading on the worrying subscale in both the U.S. and Icelandic versions.

In order to determine whether scores on the IBI were similar or different between the U.S. and Icelandic samples, it was necessary to first determine whether the two

measurement models of irrational beliefs were invariant or comparable in the first place. If scores are invariant, comparisons can continue. If they are not invariant, comparisons of U.S. and Icelandic scores may be, in Vandenberg and Lance's words, "tantamount to comparing apples and sparkplugs" (Vandenberg & Lance, 2000, p. 9). To establish whether the models were invariant or not, invariance testing was done. When invariance testing was conducted, six potential sources of differences between groups were tested: a baseline model (i.e., configural), factor loadings, intercepts, residuals, factor variances, and factor covariances. Differences were found on all six sources. More specifically, when examining the invariance of the IBI item factor loadings, intercepts, and residuals, 39 out of the IBI's 50 items were found to perform differently between the U.S. and Icelandic versions. This, in effect, rendered all but 11 items of the IBI incomparable. Significant differences were also found on the tests of factor variances and factor covariances. The 11 items that compared similarly between the U.S. and Icelandic samples are shown in Table 25, arranged by the researcher.

Table 25
Items on the IBI on which Equivalence was Tenable

Subscale	Item number and description
Worrying	16. I tend to worry about possible accidents and disasters
Worrying	q19. I often get excited or upset when things go wrong
Worrying	32. I get terribly upset and miserable when things are not the way I like them to be
Problem avoidance	q36. It is difficult for me to do unpleasant chores
Problem avoidance	q25. I shrink from facing a crisis or difficulty
Problem avoidance	q40. I often spend more time trying to think of ways of getting out of things than it would take me to do them
Rigidity	q29. One should blame oneself severely for all mistakes and wrongdoings
Rigidity	q41. Immorality should be strongly punished
Rigidity	q35. There is a right way to do everything
Need for approval	q37. It is important to me that others approve of me
Emotional irresponsibility	q30. People are disturbed not by situations but by the view they take of them

The 11 items of the IBI that were not found to be different came from all subscales, but some subscales were represented better than others. The worrying, problem avoidance, and rigidity subscales all had three items each among the invariant 11, but need for approval and emotional irresponsibility only had one item each.

Limitations of the Study

When translating and adapting an instrument to a new language and culture, most authors consider it preferable that the instrument is well-established and extensively tested in the host country (AERA, APA, & NCME, 1999; Hambleton, 2005; ITC, 2009). Little psychometric evidence was available on the IBI prior to the current study, apart from the Koopmans et al. (1994) and Bridges and Sanderman's (2002) U.S. study, and

this is arguably a limitation on the current study's investigation of the construct validity of the IBI scores. However, there does not seem to be a consensus on the best irrational beliefs measure even among REBT clinicians in the current literature, and most instruments measuring irrational beliefs seem to have insufficient or less-than-optimal psychometric properties at the time of the current study (Terjesen et al., 2009). In addition, since irrational beliefs are considered causes of psychological disturbances rather than disturbances in their own right (Ellis, 1962), they are arguably broader and more related to personality constructs than they are to specific psychological disorders, such as depression or agoraphobia (Ziegler, 1999). That said, all inventories and theories need to start development somewhere, and cross-cultural research may be a good way to discover limitations and insights that may otherwise elude researchers working with samples from only one culture. Given the fruitful outcome of the cognitive interviews in the current study, and the emerging status of REBT instruments, it may be productive to conduct extensive cognitive interviews on the nature of irrational beliefs in several countries, and use those to build the developing knowledge base of REBT theory from a culturally relative standpoint.

Another limitation to the IBI may be in the construct validity of several items on the IBI. Some items on the IBI arguably do not describe beliefs, but reactions to hypothetical situations. Terjesen et al. (2009) found this to be a fairly common situation on several instruments measuring irrationality. For example, item 16 on the IBI; "I tend to worry about possible accidents and disasters" does not describe a belief, but a worry-reaction to a hypothetical situation. Persistent worrying about hypothetical or imagined threats is irrational, but it could nevertheless be argued that overworrying, a significant

component in the IBI, is better classified as just that, and not as a belief. In contrast, items such as item 29: "One should blame oneself severely for all mistakes and wrongdoings", sound more like a description of beliefs. The difficulty with using reaction-worded items to represent beliefs is that, while development of irrational beliefs may very well be identical across cultures, the situations in which they manifest themselves may be radically different between cultures. Out of the 11 items on the IBI that were considered equal between countries in the current study, six describe reactions, and five describe beliefs (See Table 25). This begs the question of whether the IBI in its current form actually measures irrational beliefs alone, or whether it measures a combination of irrational beliefs and irrational stress responses.

Significance of Study and Recommendations for Future Research

The weak psychometric performance of the IBI may be caused by many things, perhaps the most important of which is the fact that irrational beliefs, as approached by REBT, are not a psychological disturbance in and of itself, but a personal philosophy characterized by rigid and absolutistic demands that cause the individual to become unhappy. Thus, since a personal philosophy is by design a stable pattern of verbal habits, any test measuring irrational beliefs has conceptually more in common with personality tests than it has with tests of isolated disturbances such as depression and test anxiety. If personality is an individual's tendency to respond consistently to certain circumstances, as it is specified in REBT theory (e.g., Ellis, 1962), then different cultures may, by definition, not encourage similar behaviors. REBT is well established as an applied clinical approach, but its status as a developing theory can lead to controversy around any attempts to investigate irrational beliefs through psychometric means. Still, studies like

the current study shed an important light on the limits of both REBT theory as a cross-culturally measured construct, and on the limits on attempts to measure it (Ellis, 2003; Terjesen et al., 2009). A strength of the current study is found in the abundance of findings from the translation and cognitive interview phases, which emphasize the importance of mixed methods and social-contextual approaches when psychometrically investigating challenging constructs. Perhaps, the development of REBT theory might be helped by convergent approaches to development of its instruments. In a convergent approach, as described by Werner and Campbell (1970), cross-cultural validation and development of a theory or instrument are not achieved by developing a test in one culture and then attempting to transfer it to another, but by developing the instrument in several cultures concurrently (or near-concurrently).

Traditional development of psychological inventories is by design influenced by its original host culture, but difficulties arise when a dimension exists in the target culture (e.g., Iceland) that is not touched on at all in the original culture (e.g., U.S.). However, if a dimension relevant to the construct exists in the new culture that did not exist in the original culture, the inventory is still missing a conceptual category that was missed by the original developers, undermining the universality of the instrument. For future research, cultural reviewers might be asked something akin to: "Did you find that this inventory did not touch on something that is essential to the construct in question?" If this is done, those who translate and adapt psychological tests for use in research in new cultures may find themselves discovering new constructs and variables that the original test developers' culture did not allow them to discover. It follows that the translation and adaptation of psychological instruments is potentially a two-way street, where the target

culture can actually add a new relevant dimension to an existing measure, enriching the instrument and its foundational theory alike. The original language of the IBI was Dutch, and no psychologist speaking Dutch and Icelandic could be found, so one can not make assumptions on what has been lost in translation and what has potentially been gained. The researcher did not have access to cross-validation data between the original Dutch developers and the U.S. developers, but the Dutch-U.S. translation was done in the Netherlands, by the Dutch developers, and then administered in the U.S. (K. Bridges, personal communication, 2009).

International success of a psychotherapy framework such as Rational Behavior Emotive Therapy (i.e., the basis of the Irrational Beliefs Inventory) does not mean that all parts of that framework are universally applicable; clinical work and empirical testing do not necessarily follow the same procedures, and there are several reasons why testing the cross-cultural applicability of a therapy framework can be difficult. Firstly, a therapy setting cannot by its nature be a testing ground for all facets of an extensive therapy framework. For example, a client may be suffering from overworrying, but not from an unhealthy need for others' approval. Thus, the universality of need for approval is never put to the test for that particular client. Secondly, therapists in the field will pick and choose from their skillsets depending on the needs of each client, so standardizing a therapy session in the name of research could be unethical because it might de-prioritize a client's specific needs. Thirdly, clinical concepts might not always translate even to ostensibly similar cultures as shown in the example of the term "approval" above. An exciting prospect for future research may be to look in detail at items that do *not* seem to translate across cultures. It is tenable that such items can provide just as valuable insights

as the items that actually did translate and have equivalent psychometric properties (e.g., factor loadings, item intercepts). These items might need additional probing and investigation by judgment strategies, and possibly be completely replaced with more culturally appropriate content.

A further issue in development of REBT measures is that the theory of it, and along with it the number of irrational beliefs, is constantly changing as this dissertation is written. Albert Ellis's original irrational beliefs numbered 10 (Ellis, 1962). They later became 11 (Ellis, 1997), and then he narrowed them down to three, which is the current number when this is written (Ellis & Dryden, 1987; Ellis, 2003; Terjesen et al., 2009). Koopmans et al. (1994) and Bridges and Sanderman (2002) found five factors during their analysis of the IBI. Another measure of irrational beliefs, the Smith Irrational Beliefs Inventory (SIBI), has been tested in three countries and yielded a different number of factors each time the inventory was translated: five, three, and seven, respectively (Smith, 2002). To complicate things even further for cross-cultural research on irrationality is the fact that psychological disorders are not clear-cut to begin with, and they rarely occur in isolation. For example, agoraphobia (fear of open spaces) may very well accompany an obsessive fear of germs, or a tendency to have panic attacks. As a result, any test intended to measure such a broad range of behavior samples will most likely get a few wrong by design. Nevertheless, despite the above challenges, the attempt to translate and adapt the IBI from English to Icelandic has produced some interesting findings.

As mentioned before, REBT is a theory of dispositions towards and causes for psychological disturbances, rather than a topical theory of a specific disturbance (Ziegler,

1999, 2000). As such, CFAs of measurement models underlying an instrument such as the IBI may not yield fit as good as those from an inventory measuring more discrete constructs. Unfortunately, the poor fit statistics of the IBI in both the U.S. and Iceland did not allow many cross-cultural generalizations. Yet, statistical fit indices alone do not tell the whole story. When the weak psychometric properties were interpreted in tandem with qualitative data from the translation/adaptation phase and the cognitive interviews phase, several stronger explanations of the weak fit emerged. Thus, expanding cross-cultural studies from the realm of quantitative methodology into the mixed-methods realm is likely to create insights just as powerful as might come from a well fitting CFA model, and help move the field forward as more diverse bodies of knowledge are assimilated and integrated.

A constant dilemma in all social sciences is that cultures are relative, and thoughts that are irrational in some cultures make perfect sense in others. For example, in REBT theory, the need for others' approval can only be irrational if it is an imagined threat to one's well-being, and it can manifest itself as unrealistic demands that by virtue of their unfulfillability can cause depression and narcissism alike. However, the need for approval can exist literally as a matter of life and death, and thus be rational. In some cultures, women are murdered and/or mutilated if they do not agree to marry someone selected for them in a fixed marriage, a practice referred to as "honor killings". In cultures where this custom is practiced, IBI items such as item 37, "it is important to me that others approve of me" might yield comparable mean scores as they would in the U.S., but need for approval is quite rational if one's life depends on it. Unfortunately, the practice seems to make sense to the perpetrators, and so does the victims' fear of it. In Western

cultures, one might argue that so-called "crimes of passion" can be analogues of "honor killings" in some cases, but cultures differ on the appropriate consequence for the action. In an example such as this, only a multilayered understanding of the cultural context can explain the meaning of different or similar answers on a questionnaire. A less extreme example of how need for approval does and does not translate is in the current study, where the term "approval" does not translate as approval of a person or a whole entity as such, but can translate into either agreement about something specific, such as a budget ("samthykki" in Icelandic), or acknowledgement for a specific achievement or a job well done ("vidurkenning" in Icelandic). To solve this problem in the current study, all items pertaining to need for approval needed to be translated into whole sentences that carried the meaning, but ended up sounding different to the U.S. version. For example, item 21, "I often worry about how people approve of and accept me", when translated back to English by the author from the final Icelandic version, reads: "I often worry about whether people like me and what they think of me". Translations from specific terms into sentences that convey meaning is not new in Iceland, and has often been resorted to when translating the names of psychological constructs for explanation or teaching (Friðrik H. Jónsson, personal communication, 2009). The main insight from this part of the study is that, when translating a complex operational definition of such a broad concept as irrationality into a new culture, translation and adaptation of the instrument might be a good idea, but only as a pilot to an assembly, or complete re-write of the questionnaire (van de Vijver & Poortinga, 2005).

Social contextual interpretations such as in the examples above can also apply to the substantial differences in the relative emphasis on religion between the U.S. and

Icelandic samples. People of both nations claim to be religious, yet the overwhelming majority of Icelanders do not attend church or make a lifestyle out of religion, begging the question of how such core values can be measured and compared. Questions such as these can then be extrapolated to views on matters reflected in the other factors in the IBI, that is, emotional attribution, worrying, problem avoidance, and crime and punishment. Current events can also complicate measurements. When the researcher visited Iceland to collect data in 2008, a cataclysmic economic crises had just hit Iceland. All the banks in the country had been nationalized, deep-seated corruption had been exposed, and anger at politicians and bankers was palpable in the country, as well as a profound shock to Iceland's national identity (Boyes, 2009). Since no measure of the IBI existed before the crash, it is impossible to tell whether items on the IBI pertaining to crime and punishment, such as on the rigidity scale, might have influenced the results and reflected public rage rather than irrational rage.

In general, despite mixed results from the psychometric phase, the findings of the study are helpful to the larger body of cross-cultural psychology, because even a weak instrument can provide clues about what to look for. The psychometric instrument provides a starting point, and the items, culturally equivalent or not, provide material for discussion and further analysis. The findings of the current study illustrate the already suggested need for the use of psychometric data, qualitative judgment strategies, and cognitive interviews during the translation and adaptation phase. However, getting an angle on potentially massive amounts of qualitative data from translators and others can be very difficult in a cross-cultural context. To help with that interpretation, it is recommended that a social contextual approach be used whenever possible to bridge the

gap between the cultures on one hand and the theory behind the psychometric instrument on the other. For example, if a study were to be done on the cross-cultural comparability of an irrational beliefs instrument, research could also be done on local views regarding things that such a survey would touch on (e.g., views on religion, marriage, authority, drug use, etc.). Such community-based research might prove invaluable later on, as both a guide to interpretation of cognitive interviews, and as a complement to the original theory's nomological network.

Another guide for future research is that a careful analysis of questionnaire items that function differently across groups can be just as useful for the theory as an analysis of equivalent items. Items that do *not* seem to translate across cultures can provide just as valuable insights as the items that actually did translate. The role of cognitive interviewing, qualitative judgment strategies, and local academic research into relevant variables could prove tremendously valuable at that point.

In the current study, forward-and-back-translators took on a role above and beyond what was asked of them, perhaps by design, since the forward-translation is the psychological instrument's first brush with its new adoptive language and culture. A culturally equivalent construct is one that exists in both cultures (Hambleton, 2005), and Hambleton's (2005) guidelines of cultural equivalence serve as a useful reminder for the design of any such translated instrument. However, when implementing a questionnaire such as this, it is recommended that experts in the sub-discipline of psychology in which the inventory is used be consulted for cultural equivalence. For example, if the inventory is in clinical or counseling psychology, then clinicians who use the theories in the instrument's nomological network should be consulted if possible. By the same token, if

the instrument is in Industrial/Organizational psychology, specialists in that field should be consulted whenever possible. To facilitate the translation and adaptation process, the researcher recommends that the forward-translators themselves be given a cognitive interview, in a similar fashion as the trial samples later. This might help gauge for the cultural influences on the translators themselves, and possibly help identify issues with bilingualism or even vocabulary in the translators' native language. The researcher could also write or obtain summary statements of the content domains and/or constructs intended to be measured. Such statements could be viewed on a separate page when synthesizing the items, as a construct validity check. A domain or construct checklist might also come in handy during the back-translation, since differences in back-translation might be different in meaning but not necessarily decrease the construct validity of the instrument. The cultural equivalence consultant committee could serve as a final check on the translation process, but the adaptation starts at the moment of the first forward-translation, since the translator's culture will automatically influence the translation. The role of cognitive interviews in the original language are significant, as they will give clues not only about the subjective meaning of each item, but will also provide important data on instructions for forward-translators. In addition, such qualitative data should probably be compared with comments from forward-translators, back-translators, and cultural reviewers during the process at large. To achieve this, the cognitive interview process with both a trial sample and the translators would need to be more extensive and more structured than was done in the present study. The interview structure would also need to be fairly consistent between the translation staff and the

cognitive interview sample, to facilitate the analysis and presentation of the qualitative results.

One particularly difficult aspect of translations and adaptations is the issue of cultural relevance versus cultural equivalence. Culturally relevant terms that make the instruments equivalent may not be culturally equivalent. The issue of relevance vs. equivalence is probably best decided upon by an examination of the literature upon which the inventory is based. This, in turn means that the field of cross-cultural adaptations of psychological measurement instruments relies just as much on local research and insights as it does on the original theory, regardless of the theory's culture of origin. The nature and scope of each instrument should also be taken into account. Still, since the IBI is based on a mental health counseling approach, which in and of itself lends itself to quick clinical judgments and great variability between patients, psychometric evaluation becomes tricky and subject to a selective sample of patients. What makes that selection even more difficult is that mental health issues and emotional problems may be taboo in some cultures, discussed freely in others, and simply not be salient topics in yet others. Thus, an irrational belief (e.g., "I must be thoroughly competent or I am worthless and a disgrace to myself and my family"), might be a taboo subject in cultures such as Japan, discussed openly and freely in the U.S., but irrelevant in Iceland because the term "competence" might be too general and not situation-specific enough to be relevant. However, the cognitive interviews and informal debriefings with participants in Iceland give hints towards a better instrument for REBT practitioners and personality theorists in Iceland. The nomological network of the concepts should be tied in with research in both cultures. If different versions of a psychological measurement instrument are

equivalent, care must nevertheless be taken to not enforce one culture's effect on the other. The adjustment can be reciprocal, using data from the target culture to pinpoint possible shortcomings of the original version.

Successful as REBT may be in clinical settings worldwide, its conceptual dimensions are still largely determined by clinicians (Ziegler, 1999). The extent to which this is done, however, is still unknown, because local customs and habits encoded in different languages can make the universalities hard to determine. If the REBT framework has truly captured a global human tendency, the Irrational Beliefs Inventory does not show this. Yet, the field of cross-cultural psychology is relatively new, and by examining generalities and differences across cultures, scientists can greatly improve their theories and gain precious insight into what it means to be human. It is perhaps naïve to think that cross-cultural understanding can make humans all live peacefully without conflict, but it is perfectly realistic to expect that knowledge of our man-made cultures can help applied social scientists and policymakers better identify how cultural practices have helped our species, as well as how they have been less helpful or even harmful.

In conclusion, the findings of the current study add to the growing body of literature that appears to show that development of instruments measuring irrational beliefs is still in its infancy. If the IBI itself is to be useful in the Icelandic culture, the best bet is to re-assemble rather than adapt, and it is probably advisable to survey and conduct cognitive interviews among REBT clinicians to gauge for commonly held irrational beliefs that might be local to Iceland. In fact, decentralized input on locally expressed irrational beliefs from REBT clinicians from all over the world is probably a

good idea. The process of cross-cultural adaptations of measurement instruments is time-consuming and intensive, and the construct of irrationality is ripe for input in its current developing state. As the body of research on irrationality in various cultures increases, better methods to study the relationship between a culture and the beliefs within may emerge.

References

- Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Acquadro, C., Conway, K. Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health, 11*, 509-521.
- Arnarson, T. O., Olason, D. T., Smari, J., & Sigurdsson, J. F. (2008). The Beck Depression Inventory Second Edition (BDI-II): Psychometric properties in Icelandic student and patient populations. *Nordic Journal of Psychiatry, 62*, 360-365.
- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*, 287-311.
- Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Beck, A. T., & Freeman, A. (1990). *Cognitive therapy of personality disorders*. New York: The Guilford Press.
- Boyes, R. (2009). *Meltdown Iceland*. New York: Bloomsbury.
- Bridges, K. R., & Sanderman, R., (2002) The Irrational Beliefs Inventory: Cross-cultural comparisons between American and Dutch samples. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 20*, 65-71.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guildford Press.
- Buss, D. M. (2001). Human nature and culture: An evolutionary psychological perspective. *Journal of Personality, 69*, 955-978.

- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kreammer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) Manual for Administration and Scoring*. Minneapolis: University of Minneapolis Press, 1989.
- Byrne, B. M., & van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, *10*, 107-132.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, *56*, 81-105.
- Costa P. T., Jr., & McCrae R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa P. T., Jr., McCrae R. R., & Jónsson F. H.: Validity and utility of the Revised NEO Personality Inventory: Examples from Europe. In B. DeRaad and M. Perugini (Eds.), *Big five assessment* (pp. 61-77). Göttingen, Germany: Hogrefe & Huber.
- Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.
- Danielsson, J. (2009). The first casualty of the crisis: Iceland. In Felton, A., & Reinhart, C. M. (Eds.), *The first global financial crisis of the 21st century. Part II June-December 2008*. London: Centre for Economic Policy Research (CEPR).
- Dobson, D., & Dobson, K. S. (2009). *Evidence-based practice of Cognitive-Behavioral Therapy*. New York: The Guilford Press.
- Dobson, K. S., & Dozois, D. J. (2009). Historical and philosophical bases of the cognitive-behavioral therapies. In Dobson, K. S. (Ed.), *Handbook of Cognitive Behavioral Therapies* (3rd ed.) (pp. 3-38). New York. The Guilford Press.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel Haenszel and standardization. In Holland, P. W., & Wainer, H. (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale, NJ: LEA.
- Dryden, W., David, D., & Ellis, A. (2009). Rational Emotive Behavior Therapy. In Dobson, K. S. (Ed.), *Handbook of cognitive behavioral therapies* (3rd ed.) (pp. 226-276).

- Dryden, W., DiGiuseppe, R., & Neenan, M. (2003). *A Primer on rational emotive behavior therapy* (2nd ed.). Champaign, IL: Research Press.
- Ellis, A. (1962). *Reason and emotion in psychotherapy*. New York: Lyle Stuart.
- Ellis, A., & Harper, R. A. (1997). *A guide to rational living* (3rd ed.). Chatsworth, CA: Wilshire.
- Ellis, A., & Dryden, W. (1987). *The practice of rational emotive therapy (RET)*. New York: Springer.
- Ellis, A., (2003a). The relationship of rational emotive behavior therapy (REBT) to social psychology. *Journal of Rational-Emotive Cognitive-Behavior Therapy*, 21, 5-20.
- Ellis, A. (2003b). Reasons why rational emotive behavior therapy is relatively neglected in the professional and scientific literature. *Journal of Rational-Emotive Cognitive-Behavior Therapy*, 21, 245-253.
- Ellis, A., & Dryden, W. (2007). *The practice of rational emotive behavior therapy* (2nd ed.). New York: Springer.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.
- Eysenck, H. J., & Eysenck, S. B. G. (1983). Recent advances in the cross-cultural study of personality. In Butcher, N., & Spielberger, C. D. (Eds.), *Advances in personality assessment. Volume 2* (pp. 41-70). Hillsdale, NJ: LEA.
- Ercikan, K., & Koh, K. (2005). Examining the construct comparability of the English and French versions of TIMSS. *International Journal of Testing*, 5, 23-35.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286-299.
- Furlan, A. S., Cassidy, J. C., & Pérez, E. R. (2009). Adapting the Cognitive Test Anxiety Scale for use with Argentinean university students. *International Journal of Testing*, 9, 3-19.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the interpretation of normative assessment instruments. *Psychological Assessment*, 6, 304-312.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186-192.

- Guðmundsson, E. (2005). Þýðing og staðfærsla sálfræðilegra prófa [Translations and adaptations of psychological tests]. *Sálfræðiritið*, 10-11, 23-40.
- Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33, 587–606.
- Goldberg, D. P., & Williams, P. (1988). *A user's guide to the General Health Questionnaire*. Berkshire: NFER-NELSON.
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, and C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: LEA.
- Hambleton, R. K. (2006). Good practices for identifying differential item functioning. *Medical Care*, 44, 182-188.
- Haraldsson, E., Smári, J., & Gylfason, H. F. (2001). MMPI-2: Athugun á íslenskri gerð prófsins [MMPI-2: An investigation of its Icelandic version]. *Sálfræðiritið*, 7, 51-65.
- Harkness, J., Mohler, P. P., & van de Vijver, F. J. R., (Eds.) (2003). Comparative research: *Cross-cultural survey methods*. Hoboken, NJ: Wiley.
- He, W., & Wolfe, E. W. (2010). Item equivalence in English and Chinese translation of a cognitive development test for preschoolers. *International Journal of Testing*, 10, 80- 94.
- Himle, D. P., Hnat, S., Thyer, B., & Papsdorf, J. D. (1985). Factor structure of the Rational Behaviour Inventory. *Journal of Clinical Psychology*, 41, 368-371.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer, & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale: LEA.
- Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, 14, 332-346.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus alternatives. *Structural Equation Modeling*, 6, 1-55.

- Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology*, *16*, 131-152.
- International Test Commission (2001). International guidelines for test use. *International Journal of Testing*, *1*, 93-114.
- Jones, R. G. (1968). A factored measure of Ellis' irrational belief system with personality, maladjustment correlates (Doctoral Dissertation, Texas Tech University, 1969). *Dissertations Abstracts International B*, *29*, 4379.
- Jónsdóttir, S. M., Þorsteinsdóttir, G., & Smári, J. (2005). Próffræðilegir eiginleikar íslenskrar gerðar Bulimia Test-Revised (BULIT-R) prófsins [Psychometric properties of the Icelandic version of the Bulimia Test-Revised (BULIT-R)]. *Læknablaðið*, *91*, 923-928.
- Jónsson, F. H., & Bergþórsson, A. (2004). Fyrstu niðurstöður úr stöðlun NEO-PI-R á Íslandi [First results from the standardization of the NEO-PI-R in Iceland]. *Sálfræðiritið*, *9*, 9-16.
- Jónsson, F. H. (2005). Próffræðilegir eiginleikar íslenskrar útgáfu NEO-FFIR [Psychometric properties of the Icelandic version of the NEO-FFIR]. In Úlfar Hauksson (Ed.), *Rannsóknir í Félagsvísindum*, *6* (429-439). Reykjavík: Félagsvísindastofnun Háskóla Íslands.
- Koller, M., Aaronson, N. K., Blazeby, J., Bottomley, A., Dewolf, L., Fayers, P., Johnson, C., Ramage, J., Scott, N., & West, K. (2007). Translation procedures for standardised quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach. *European Journal of Cancer*, *43*, 1810-1820.
- Koopmans, P. C., Sanderman, R., Timmerman, I., & Emmelkamp, P. M. G. (1994). The Irrational Beliefs Inventory (IBI): Development and psychometric evaluation. *European Journal of Psychological Assessment*, *10*, 15-27.
- Köbben, A., J. F. (1970). Cause and Intention. In R. Naroll, and R. Cohen (Eds.), *A handbook of method in cultural anthropology* (pp. 89-98). Garden City, NY: Natural History Press.
- Levin, K., Willis, G. B., Forsyth, B. H., Norberg, A., Stapleton, Kudela, M., Stark, D., & Thompson, F. E. (2009). Using cognitive interviews to evaluate the Spanish-language translation of a dietary questionnaire. *Survey Research Methods*, *3*, 13-25.
- Lohr, J. M., & Bonge, D. (1982). The factorial validity of the Irrational Beliefs Test: A psychometric investigation. *Cognitive Therapy and Research*, *6*, 225-230.

- Lord, F. M. (1977). A study of item bias, using item characteristic curve theory. In Y. H. Poortinga (Ed.), *Basic problems in cross-cultural psychology*. Amsterdam: Swets & Zeitlinger.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Lord, F. M., (1980). *Applications of item response theory to practical testing problems*. Hillsdale: LEA.
- Malcarne, V., L., Fernandez, S., & Flores, L. (2005). Factorial validity of the Multidimensional Health Locus of Control Scales for three American ethnic groups. *American Journal of Health Psychology, 10*, 657-667.
- Malouff, J. (2009). Commentary on the current status of assessment in Rational-Emotive and Cognitive-Behavior Therapy. *Journal of Rational-Emotive and Cognitive-Behavior Therapy, 27*, 136-140.
- McCrae, R. R. (2001). Trait psychology and culture: Exploring intercultural comparisons. *Journal of Personality, 69*, 819-846.
- McDermut, W., & Haaga, D. A. F. (2009). Assessment and Diagnostic Issues in Rational Emotive Behavior Therapy: Introduction to the Special Issue. *Journal of Rational-Emotive Cognitive-Behavior Therapy, 27*, 79-82.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525-543.
- Messick, S. M. (1988). The once and future issues of validity. Assessing the meaning and consequences of measurement. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 33-45). Hillsdale, NJ: LEA.
- Ministry of Education, Science, and Culture (2002). *The Educational System in Iceland*. Reykjavik: Author.
- Morris, D. (1969). *The human zoo. A zoologist's classic study of the human animal*. New York: Kodansha Globe.
- Muthén, B., & Muthén, L. (2008). *MPlus 3.0 (computer program)*. Los Angeles: Muthén & Muthén.
- OECD (2010a). *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I)*. <http://dx.doi.org/10.1787/9789264091450-en>

- OECD (2010b). *PISA 2009 Results: Overcoming social background – Equity in Learning Opportunities and Backgrounds (Volume II)*. <http://dx.doi.org/10.1787/9789264091504-en>
- OECD (2010c). *PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V)*. <http://dx.doi.org/10.1787/9789264091580-en>
- Pan, Y. (2004) *Cognitive interviews in languages other than English: Methodological and research issues*. Paper Presented at the 59th Annual Conference of the American Association for Public Opinion Research, Phoenix, AZ.
- Poortinga, Y. H., Bijnen, E. J., & Hagenaars, J. A. (1994). Identifying bias in cross-cultural data. In M. A. Croon & F. J. R. van de Vijver (Eds.), *Viability of mathematical models in the social behavioral sciences*. Lisse: Swets & Zeitlinger.
- Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 3, 369-379.
- Sanderman, R., Mersch, P-P., van der Sleen, J., Emmelkamp, P. M. G., & Ormel, J. (1987). The Rational Behavior Inventory (RBI): A psychometric evaluation. *Personality & Individual Differences*, 8, 561-569.
- Sevigny, S., Savard, D. & Beaudoin, I. (2009). Comparability of writing assessment scores across languages: Searching for evidence of valid interpretations. *International Journal of Testing*, 9, 134–150.
- Scott, N. W., Fayers, P. M., Aaronson, N. K., Bottomley, A., de Graeff, A., Groenvold, M., Gundy, C., Koller, M., Petersen, M. A., & Sprangers, M. A. G. (2010). Differential item functioning (DIF) analyses of health-related quality of life instruments using logistic regression. *Health and Quality of Life Outcomes*, 8 : 81. Retrieved February 12, 2011, from <http://www.hqlo.com/content/8/1/81>
- Shorkey, C. T., & Whiteman, V. L. (1977). Development of the Rational Behavior Inventory: Initial validity and reliability. *Educational and Psychological Measurement*, 37, 527-534.
- Smith, J. C., Rausch, S. M., & Jenks-Ketterman, J. D (2004). Factor structure of the Smith Irrational Beliefs Inventory: Results of the analyses on six independent samples. *Psychological Reports*, 95, 696-704.
- Sireci, S.G., Yang, Y., Harter, J., & Ehrlich, E.J. (2006). Evaluating guidelines for test adaptations. *Journal of Cross-Cultural Psychology*, 37, 557-567.
- Skinner, B. F. (1971). *Beyond freedom and dignity*. Indianapolis: Hackett.

- Solano-Flores, G., Backhoff, E., & Contreras-Niño, L. Á. (2009). Theory of test translation error. *International Journal of Testing*, 9, 78-91.
- Spielberger, C. D., Moscoso, M. S., & Brunner, T. M. (2005). Cross-cultural assessment of emotional states and personality traits. In R. K. Hambleton, P. F. Merenda, and C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Erlbaum.
- SPSS for Windows, Rel. 18. 2010. Chicago: SPSS Inc.
- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, 91, 1292–1306.
- Statistics Iceland (2011). *Statistical Table – Population – Overview*. Retrieved February 1, 2011, from <http://www.statice.is>.
- Steenkamp, J-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *The Journal of Consumer Research*, 25, 78-90.
- Teresi, J. A., Ramirez, M., Lai, J-S., & Silver, S. (2008). Occurrences and sources of Differential item functioning (DIF) in patient-reported outcome measures: Description of DIF methods, and review of measures of depression, quality of life and general health. *Psychology Science Quarterly*, 50, 538-612.
- Terjesen, M., D., Salhany, J., & Sciutto, M. J. (2009). A Psychometric Review of Measures of Irrational Beliefs: Implications for Psychotherapy. *Journal of Rational-Emotive Cognitive-Behavior Therapy*, 27, 83–96.
- Thelen, M. H., Mintz, L. B, Vander Wal, J. S. (1996). The Bulimia Test-Revised: Validation with DSM-IV criteria for bulimia nervosa. *Psychological Assessment*, 8, 219-21.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis. Understanding concepts and applications*. Washington, D.C.: American Psychological Association.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.
- Torp, A. (2004). Norræn tungumál að fornu og nýju [Nordic languages in ancient and modern times]. In Sletten, I. S. (Ed.), *Norðurlandamálin með rótum og fótum*. Nord 2004:010. Copenhagen, Denmark: Nordic Council of Ministers' Secretariat.
- International Association for the Evaluation of Educational Assessment (IEA) (2008). *Trends in International Mathematics and Science Study*. Retrieved November 6, 2008, from <http://timss.bc.edu>

- United Nations Educational, Scientific and Cultural Organization (2002). *UNESCO Universal Declaration on Cultural Diversity*. Retrieved July 14, 2008, from http://www.unesco.org/education/imld_2002/universal_decla.shtml
- United Nations High Commissioner for Refugees (2010). *Freedom in the World 2010 - Iceland*. Retrieved February 16, 2011, from <http://www.unhcr.org/refworld/docid/4c0ceae5.html>
- United Nations Human Development Programme (2009). *Human Development Report 2009. Overcoming barriers: Human mobility and development*. Retrieved February 4, 2011, from <http://hdr.undp.org/en/reports/global/hdr2009/chapters/>
- United Nations Human Development Programme (2010). *Human Development Report 2010 - 20th Anniversary Edition. The Real Wealth of Nations: Pathways to Human Development*. Retrieved February 4, 2011, from <http://hdr.undp.org/en/reports/global/hdr2010/chapters/>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement variance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-70.
- Van de Vijver, F. J. R., & Poortinga, Y. (2005). Conceptual and methodological issues in adapting tests. In Hambleton, R., Merenda, P. F., & Spielberger, C. D. (Eds.), *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. (pp. 39-63). Mahwah: LEA.
- Van de Vijver, F. J. R. (2008). Personality assessment of global talent: Conceptual and methodological issues. *International Journal of Testing, 8*, 304-314.
- Werner, O., & Campbell, D. T. (1970). Translating, working through interpreters, and the problem of decentering. In R. Naroll and R. Cohen (Eds.), *A handbook of cultural anthropology* (pp. 398-419). Garden City: The Natural History Press.
- Wicherts, J. M., & Dolan, C. V. (in press). Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities. *Educational Measurement: Issues & Practice*.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health, 8*, 94-104.
- World Health Organization (2000). *Highlights on Health in Iceland*. Copenhagen, Denmark: WHO

- World Values Survey (2005). Official Data File, v.20090901, 2009. *World Values Survey Association* (www.worldvaluessurvey.org).
- Wu., A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*, 1-26.
- Ziegler, D. J. (1999). The construct of personality in rational emotive behavior therapy (REBT). *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 17*, 19-32.
- Ziegler, D. J. (2000). Basic assumptions concerning human nature underlying rational emotive behavior theory (REBT) personality theory. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 18*, 67-85.

Appendices

Appendix A

The Irrational Beliefs Inventory - U.S. Version Arranged by Subscales

Answering options - Five-point Likert scale

1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

Worrying

1. If I can't keep something from happening, I don't worry about it.
2. I worry a lot about certain things in the future.
6. I often can't get my mind off some concern.
7. I tend to become terribly upset when things are not the way I would like them to be.
8. I am fairly easygoing about life.
10. I hardly ever think of such things as death or atomic war.
16. I tend to worry about possible accidents and disasters.
19. I often get excited or upset when things go wrong.
22. Sometimes I can't get a fear off my mind.
26. I feel little anxiety over unexpected danger or future events.
28. Frustrations upset me.
32. I get terribly upset and miserable when things are not the way I like them to be.

Rigidity

3. Certain people are bad or wicked and should be severely punished for their sins.
4. People should observe moral laws more strictly than they do.
9. Punishing oneself for all errors will prevent future mistakes.
14. Those who do wrong deserve to be blamed.
18. A large number of people are guilty of bad sexual conduct.
20. It is sinful to doubt the Bible.
24. The fear of punishment helps people to be good.
29. One should blame oneself severely for all mistakes and wrongdoings.
33. More people should face up to the unpleasantness of life.
34. Helping others is the very basis of life.
35. There is a right way to do everything.
38. Too many evil persons escape the punishment they deserve.
39. It is realistic to expect that there should be no incompatibility in marriage.
41. Immorality should be strongly punished.

Appendix A Continued

Problem avoidance

- 11. I avoid facing my problems.
- 13. I usually try to avoid chores which I dislike doing.
- 25. I shrink from facing a crisis or difficulty.
- 27. If something is necessary, I do it even if it is unpleasant.
- 31. I usually put off important decisions.
- 36. It is difficult for me to do unpleasant chores.
- 40. I often spend more time trying to think of ways of getting out of things than it would take me to do them.
- 44. One should rebel against doing unpleasant things, however necessary, if doing them is unpleasant.
- 45. I can't stand to take chances.
- 47. I dislike responsibility.

Need for approval

- 5. I want everyone to like me.
- 21. I often worry about how people approve of and accept me.
- 23. I hate to fail at anything.
- 37. It is important to me that others approve of me.
- 43. What others think of you is most important.
- 48. Although I like approval, it's not a real need for me.
- 50. I have considerable concern with what people are feeling about me.

Emotional irresponsibility

- 12. A person won't stay angry or blue long, unless he keeps himself that way.
- 15. If a person wants to, he can be happy under almost any circumstances.
- 17. Nothing is upsetting in itself - only in the way you interpret it.
- 30. People are disturbed not by situations but by the view they take of them.
- 42. There is never any reason to remain sorrowful for very long.
- 46. Man makes his own hell within himself.
- 49. People who are miserable have usually made themselves that way.

Summary of Scales

- Worrying 1,2,6,7,8,10,16,19,22,26,28,32
- Rigidity 3,4,9,14,18,20,24,29,33,34,35,38,39,41
- Problem avoidance 11,13,25,27,31,36,40,44,45,47
- Demand for approval 5,21,23,37,43,48,50
- Emotional irresponsibility 12,15,17,30,42,46,49

Scoring Rules

The following items are phrased in a rational direction and should be transformed in the opposite direction, by recoding (1=5, 2=4, 3=3, 4=2, 5=1):

- 1, 8, 10, 12, 15, 17, 26, 27, 30, 42, 46, 48, 49

Appendix B

The Irrational Beliefs Inventory - Icelandic Version Arranged by Subscales

Svarmöguleikar á fimm punkta Likert-kvarða

1. Mjög ósammála
2. Ósammála
3. Hlutlaus
4. Sammála
5. Mjög sammála

Áhyggjur

1. Ef ég get ekki komið í veg fyrir að eitthvað gerist, þá hef ég ekki áhyggjur af því.
2. Ég hef þungar áhyggjur af hlutum sem gætu komið upp á í framtíðinni
6. Þegar ég hef áhyggjur af einhverju tilteknu á ég oft erfitt með að leiða hugann frá því.
7. Þegar eitthvað er ekki alveg eins og ég vil hafa það, hef ég tilhneigingu til að láta það fara alltof mikið í skapið á mér.
8. Ég er frekar afslappaður/afslöppuð yfir lífinu almennt.
10. Ég hugsa sjaldan eða aldrei um hluti eins og dauðann, kjarnorkustríð, hryðjuverk og þess háttar.
16. Ég hef tilhneigingu til að hafa áhyggjur af slysum og meiriháttar áföllum sem framtíðin gæti borið í skauti sér.
19. Ég verð oft æst(ur) eða kemst í uppnám þegar eitthvað fer úrskeiðis.
22. Mér gengur stundum erfiðlega að hreinsa hugann þegar ég hef áhyggjur eða kvíði einhverju.
26. Ég kvíði lítið sem ekkert fyrir framtíðinni.
28. Ég pirrast og verð jafnvel æst(ur) þegar ég lendi í erfiðleikum með eitthvað.
32. Ég verð afskaplega æst(ur) og mér líður ömurlega þegar hlutirnir eru ekki eins og ég vil hafa þá.

Ósveigjanleiki

3. Sumt fólk er vont og illgjarnt að eðlisfari og það verðskuldar þunga refsingu fyrir misgjörðir sínar.
4. Fólk ætti að virða almenn siðabod og siðferðisleg gildi betur en það gerir.
9. Ef maður refsar sjálfum sér nógu harðlega fyrir allar misgjörðir er hægt að forðast mistök í framtíðinni.
14. Fólk sem gerir illt á allt vont skilið.
18. Margt fólk hegðar sér illa í kynferðismálum.
20. Það er syndsamlegt að efast um það sem stendur í Biblíunni..
24. Hræðslan við refsingu heldur fólki heiðarlegu.
29. Maður á að taka harkalega í lurginn á sjálfum sér fyrir eigin mistök og misgjörðir.
33. Fleira fólk þarf að horfast í augu við að lífið er stundum erfitt.
34. Að hjálpa öðrum er það sem gefur lífinu gildi..
35. Það er til rétt aðferð við að gera hvað sem er.
38. Of mörg illmenni sleppa við refsinguna sem þau verðskulda.
39. Það er raunhæft að búast við misfellulausu hjónabandi.
41. Það ættu að vera strangar refsingar við hvers konar siðleysi.

Continued on next page

Appendix B Continued

Hliðrun

11. Ég forðast að horfast í augu við vandamál mín.
13. Ég reyni yfirleitt að koma mér undan skylduverkum sem mér þykja leiðinleg.
25. Ég forðast að fást við aðsteðjandi erfiðleika eða vandamál.
27. Ef ég þarf að gera eitthvað, þá geri ég það, jafnvel þó það sé óþægilegt.
31. Ég fresta yfirleitt mikilvægum ákvörðunum.
36. Ég á erfitt með að gera leiðinleg skylduverk.
40. Ég eyði oft meiri tíma í að komast hjá því að gera hluti heldur en það tæki mig að framkvæma þá.
44. Maður ætti að streitast á móti því að vinna leiðinleg og óþægileg verk, alveg sama hversu nauðsynleg þau eru.
45. Mér er meinilla við að taka hvers konar áhættu.
47. Mér finnst vont að þurfa að bera ábyrgð.

Þörf fyrir viðurkenningu

5. Ég vil að öllum líki vel við mig.
21. Ég hef oft áhyggjur af því hvort fólki líki vel við mig og hvaða álit það hefur á mér..
23. Ég þoli alls ekki að mistakast eitthvað.
37. Það er mér mikilvægt að aðrir viðurkenni mig.
43. Álit annarra á mér er mjög mikilvægt.
48. Þótt mér þyki gott að falla öðrum í geð, þá er það mér samt ekki nauðsynlegt.
50. Ég hef talsverðar áhyggjur af því hvað fólki finnst um mig.

Tilfinningalegt raunsæi

12. Fólk er almennt ekki reitt eða einmana mjög lengi nema það velti sér uppúr eigin eynd.
15. Ef viljinn er fyrir hendi er hægt að vera hamingjusamur við nánast hvaða kringumstæður sem er.
17. Ekkert er í sjálfu sér ógnvekjandi eða skelfilegt - þetta fer allt eftir því hvernig maður túlkar hlutina.
30. Túlkun fólks á atburðum og aðstæðum kemur því í uppnám, ekki atburðirnir sjálfir.
42. Það er aldrei ástæða til að vera sorgmædd(ur) mjög lengi.
46. Fólk skapar sjálfu sér eigin eynd og vanlíðan
49. Fólk sem líður alltaf mjög illa getur yfirleitt sjálfu sér um kennt.

Athugasemd: Eftirfarandi atriði eru orðuð á raunsæjan hátt og þurfa því að vera kóðuð öfugt (1=5, 2=4, 3=3, 4=2, 5=1) áður en til úrvinnslu kemur:

1, 8, 10, 12, 15, 17, 26, 27, 30, 42, 46, 48, 4

Appendix C

Recruitment notes to professors - U.S. and Icelandic Versions

Dear professor [name here],

My name is Gudmundur Heimisson, and I am a doctoral candidate in Curriculum & Instruction, with emphasis on cross-cultural psychology and measurement at the College of Education at The University of South Florida. I am in the process of collecting data for my dissertation, and need participants between the ages of 20-64 years old. I am writing this to request permission to collect data in your classroom. The data-collection is in the form of a psychological inventory called the Irrational Beliefs Inventory, which takes approximately 15 minutes to fill out. All answers are anonymous. This study has been approved by the USF Institutional Review Board, IRB # 107751- G .

Your help would be important, and dearly appreciated. If you can help, please let me know which times would be good for you.

Best regards,

Gudmundur T. Heimisson

Email: gheimiss@usf.edu

Cell: 813-205-3674

Kæri [nafn]

Ég heiti Guðmundur Torfi Heimisson og er doktorsnemi við kennsluvísindadeild University of South Florida í Tampa með áherslu á samanburði sálfræðilegra mælingaraðferða milli menningarsvæða. Um þessar mundir er ég að safna gögnum fyrir doktorsverkefni mitt og þarfnast þátttakenda á aldrinum 20-64 ára. Mér þætti vænt um að fá að koma í tíma til þín og leggja fyrir spurningalistann Irrational Beliefs Inventory (íslenskt nafn á listann er enn í vinnslu), sem byggt er á raunsæis- og tilfinningatengdri atferlismeðferð (Rational Emotive Behavior Therapy). Heildartími fyrir lagnar er um 15 mínútur að hámarki. Öll svör þátttakenda eru nafnlaus og meðferð gagna verður samkvæmt lögum um meðferð persónuupplýsinga. Rannsóknin er á skrá hjá Persónuvernd.

Ef þetta er hægt, segðu mér hvaða tími hentar þér best. Hjálp þín væri mér mikils virði.

Bestu kveðjur,

Guðmundur Torfi

Appendix D

Informed Consent Note to Participants - U.S. and Icelandic Versions

This is an international research project, done for a doctoral dissertation, by Guðmundur Heimisson, a doctoral student of Educational Measurement and Research at the University of South Florida. The purpose of this research project is to translate a psychological survey, the Irrational Beliefs Inventory, from English to Icelandic, and adapt it to the Icelandic culture. Data will be collected both in the United States (about 1,000 participants) and Iceland (about 1,000 participants).

The main focus of this study is not on the beliefs or feelings of participants, but on the psychometric properties of the Irrational Beliefs Inventory. You may skip any or all questions on the survey, and you may leave the study at any point that you wish, but this study is completely anonymous, and your answers cannot be traceable back to you.

Due to the anonymous nature of the study, it will not be possible to provide individualized results (such as a score on the inventory). However, if you would like, you can contact me for general information about the results of the study:

Guðmundur T. Heimisson
email: heimisso@coedu.usf.edu

By removing this sheet from the test booklet and filling out the questionnaire, you give your informed consent for the data you submit to be used in research at the University of South Florida.

Kæri þátttakandi:

Það sem hér fer á eftir er alþjóðlegt rannsóknarverkefni, unnið við University of South Florida í samvinnu við nokkrar íslenskar menntastofnanir. Umsjón með verkefninu hefur Guðmundur Torfi Heimisson, doktorsnemi í aðferðafræði og sálfræði við USF. Tilgangur verkefnisins er að laga bandarískan spurningalista, Irrational Beliefs Inventory, að íslenskri tungu og menningu. Listinn hefur reynst vel í BNA, Hollandi og fleiri löndum sem hjálpartæki við greiningu á margs konar vandamálum. Gögnum verður aflað í Bandaríkjunum og á Íslandi og um það bil 1.000 manns í hvoru landi munu taka þátt.

Þú mátt sleppa eins mörgum spurningum og þú vilt og hætta þátttöku hvenær sem er. Þú nýtur nafnleyndar við þátttökuna og ekki verður hægt að rekja svör til einstakra þátttakenda. Þar að auki má geta þess að áherslan í stöðlunarverkefnum sem þessum er á próffræðilegum eiginleikum spurningalistanna, ekki einstökum svörum þátttakenda.

Vegna nafnleyndar verður ekki hægt að gefa þátttakendum eigin niðurstöður en ef þú vilt fræðast um rannsóknina almennt geturðu haft samband við mig í tölvupósti:

Guðmundur Torfi Heimisson.
Tölvupóstur: heimisso@coedu.usf.edu

Appendix E

IBI - Survey/checklist for translators, back-translators, and cultural validators - U.S. and Icelandic Versions

[Forms condensed]

U.S. version:

1. Name: _____
2. How long did the work take you: Approx. _____ hours.
3. Did you do this all in one session, or in several sessions. If several, please explain:
4. Did any items raise flags because of strange/archaic/culturally inappropriate wording or that you would not expect to see in a psychological test?
Please state item # and concern below:
5. Did you encounter any items that were particularly difficult to translate, and if so, why?¹
Please state item # and reason below:
6. Any other thoughts/concerns/difficulties/comments/:

Icelandic version:

1. Nafn: _____
2. Hversu lengi tók verkið þig: U.þ.b. _____ klst.
3. Vannstu verkið allt í einni lotu eða fleirum? Ef fleirum, hvers vegna?
4. Stungu einhver atriði í augu vegna einkennilegs eða óþjáls orðalags? Hvers vegna?
Ef svo, vinsamlegast taktu þau til hér fyrir neðan:
5. Voru einhver atriði erfiðari en önnur? Ef svo, hvers vegna?
Vinsamlegast taktu til atriðið og ástæðuna hér fyrir neðan:
6. Aðrar athugasemdir:

¹ In the Icelandic version of the survey, the words for "to translate" were omitted, which allowed the form to be used by the Icelandic cultural validators as well as the translators.