BRITISH COUNCIL

BRITISH COUNCIL MONOGRAPHS ON MODERN LANGUAGE TESTING 3

Series Editors: Barry O'Sullivan and Karen Dunn

Validity

Theoretical Development and Integrated Arguments

Micheline Chalhoub-Deville
Barry O'Sullivan

VALIDITY

Theoretical Development and Integrated Arguments

Micheline Chalhoub-Deville and Barry O'Sullivan

equinox

SHEFFIELD UK BRISTOL CT

CONTENTS

List of figures		viii
List of tables		ix
Chapter 1	Introduction	1
Chapter 2	Historical overview of validity	13
Chapter 3	Principled design, test development and validation	59
Chapter 4	Validity and consequences	128
Chapter 5	An integrated argument-based approach to validation	146
Chapter 6	Conclusion: Validation, localisation and case studies	158
References		168
Appendix		189
Index		201

LIST OF FIGURES

Figure 1: The speaking test validation framework (O'Sullivan and Weir, 2002) 22

Figure 2: Kane's measurement procedure and interpretive argument for trait interpretations (2006) 47

Figure 3: Kane's (2006) IA approach adapted by Aryadoust (2009: 1193; 2013) 50

Figure 4: Inferential links from consequences to assessment performance (Bachman & Palmer, 2010) 61

Figure 5: The structure of the technical evaluation of educational testing (Lissitz & Samuelsen, 2007a) 70

Figure 6: Schematic representation of test design evidentiary structure (Mislevy, Steinberg & Almond, 2003) 79

Figure 7: A universal validity system (Embretson, 2008) 84

Figure 8: Revised test validation model (O'Sullivan, 2014) 94

Figure 9: Khalifa & Weir's model of reading progression (2009) 117

Figure 10: Role and responsibility allocation for consequences in validity (Chalhoub-Deville, 2016) 144

Figure 11: Validation model as integrated arguments (Chalhoub-Deville course materials, 2012–present) 149

Figure 12: Socio-cognitive model as integrated arguments 155

LIST OF TABLES

Table 1:	Definitions of content, criterion and construct validity	7
Table 2:	Methodologies associated with content, criterion and construct validity	8
Table 3:	Summary of significant influences on UK test development (based on Weir et al., 2013)	16
Table 4:	Messick's progressive matrix (1989)	31
Table 5:	Weideman's matrix (2012)	34
Table 6:	McNamara & Roever's matrix (2006)	35
Table 7:	Definition of validity in the *Standards* (1999 and 2014)	36
Table 8:	Validity as a unitary concept in the *Standards* (1999 and 2014)	37
Table 9:	Construct validity in the *Standards* (1999 and 2014)	39
Table 10:	Sources of validity evidence with some examples (from the *Standards* of 1999 and 2014)	43
Table 11:	Formulation of validity research in the *Standards* of 1999 and 2014	44
Table 12:	Nature of arguments (Kane, 2006)	53
Table 13:	Attending to stakeholder groups in Bachman & Palmer's AUA (2010)	66
Table 14:	Internal factors that should be considered for the systematic evaluation of content validity (Lissitz & Samuelsen 2007a)	72
Table 15:	The UVS test development and operational segment	84
Table 16:	The UVS meaning-oriented segment	85
Table 17:	The UVS significance-directed segment	86

Table 18: Comparison of some validity issues in the UK and the US 92

Table 19: Cognitive dimension of Aptis Reading Task 4 (O'Sullivan & Dunlea, 2015) 103

Table 20: Linking Messick's six validation criteria to the socio-cognitive model 124

Table 21: Conceptualising consequences within validity in reform-driven testing 141

Table A1: Levels of localisation in the Aptis test system (O'Sullivan & Dunlea, 2015) 191

Table A2: Overview of development activities, British Council / TUFS Project 197