Revista Interamericana de Psicología

Interamerican Journal of Psychology

Vol. 21, Nos. 1 & 2, pp. i-xv, 1-107 1987

DIRECTOR/EDITOR

Luis M. Laosa

Educational Testing Service, Princeton, New Jersey, U. S. A.

ISSN: 0034-9690

La *Revista Interamericana de Psicología/Interamerican Journal of Psychology* es la publicación oficial de la Sociedad Interamericana de Psicología/Interamerican Society of Psychology. El propósito de la *Revista* es servir como un medio de intercambio de información científica y profesional entre psicólogos, a través del Hemisferio Occidental (América del Norte, Central y del Sur y el Caribe). La *Revista* está diseñada para facilitar la comunicación, promover un mayor entendimiento y fomentar la cooperación entre los científicos del comportamiento, a través de las Américas. Su objetivo es interesar a todos los psicólogos en el Hemisferio Occidental. Con este fin, la *Revista* acepta manuscritos en todas las áreas de psicología general y las disciplinas relacionadas. Se aceptan los manuscritos en inglés, español y portugués.

Los manuscritos aceptados caen dentro de varias categorías: informes sobre estudios empíricos originales, artículos teóricos, revisiones integrativas o críticas de literatura y contribuciones metodológicas. Ocasionalmente, la *Revista* publica también artículos que tratan sobre temas importantes de política relacionados a la psicología como una ciencia y/o una profesión. Aunque está más orientada hacia la investigación básica y la teoría, la *Revista* publica también artículos que tratan más directamente con problemas de tecnología y de aplicación. Trabajos enfocados al contexto cultural del comportamiento y desarrollo humanos son especialmente bien recibidos, pero con seguridad el alcance de los manuscritos aceptados no se limita a ese enfoque.

Dirección de correo del Director: Luis M. Laosa, Educational Testing Service, Princeton, New Jersey 08541, U.S.A.

El Director envía cada manuscrito como mínimo a dos consultores, con frecuencia a más de dos, en el proceso de crítica y evaluación. Indiferentemente a la categoría a la que pertenecen, el criterio fundamental usado en la crítica y evaluación de los manuscritos es la excelencia científica y académica.

Se publican los manuscritos en la misma lengua en que se presentan. Las pruebas de galerada, junto con el manuscrito editado para imprimir, son enviados al autor principal con el objeto de que el mismo autor pueda controlar las pruebas. Es una regla de esta revista el prohibir a los autores la presentación de manuscritos que se publican en otras partes; también se prohibe la presentación simultánea de un manuscrito a dos o más revistas u otras publicaciones.

Reseñas de libros. Las reseñas de libros solicitadas y los ensayos sobre libros se publican en la Sección de Reseñas de Libros. Editores de Reseñas de Libros: Rubén Ardila, Editor de Reseñas de Libros en Español, Universidad Nacional de Colombia, Bogotá, Colombia. Jon D. Swartz, Editor de Reseñas de Libros en Inglés, Southwestern University, Georgetown, Texas, U.S.A.

Suscripciones. La *Revista Interamericana de Psicología/Interamerican Journal of Psychology* se envía por correo a todos los miembros de la Sociedad Interamericana de Psicología/Interamerican Society of Psychology (SIP/ISP). Todas las personas cuyo trabajo profesional, escolástico o científico implica un interés relacionado con cualquier aspecto de investigación o práctica psicológica son elegibles como miembros. Envíe toda correspondencia acerca de suscripciones a una de las siguientes

direcciones: SIP/ISP Suscripciones Para Individuos, Departamento de Psicología, Universidad de Puerto Rico, Apartado 23174, Estación UPR, Río Piedras, Puerto Rico 00931-3174. SIP/ISP Suscripciones Para Instituciones, 124 Professional Center Parkway, San Rafael, California 94903, U.S.A. También están disponibles los números anteriores de los volúmenes desde estas mismas direcciones.

The *Revista Interamericana de Psicología/Interamerican Journal of Psychology* is the official journal of the Sociedad Interamericana de Psicología/Interamerican Society of Psychology. The purpose of the *Journal* is to serve as a medium to exchange scientific and professional information among psychologists throughout the Western Hemisphere (North, Central, and South America and the Caribbean). The *Journal* is designed to facilitate communication, promote greater understanding, and foster cooperation among behavioral scientists across the Americas. It is intended to be of interest to all psychologists in the Western Hemisphere. To these ends, the *Journal* accepts manuscripts in all areas of general psychology and related disciplines. Manuscripts are accepted in English, Spanish, or Portuguese.

Accepted manuscripts fall within several categories: reports of original empirical studies, theoretical articles, integrative or critical literature reviews, and methodological contributions. Occasionally, the *Journal* also publishes articles dealing with significant policy issues related to psychology as a science and/or a profession. Although it is more oriented to basic research and theory, the *Journal* also publishes articles dealing more directly with technological and applied problems. Research focusing on the cultural context of human behavior and development are especially welcome, but assuredly the scope of accepted manuscripts is not limited to this focus.

Editor's address: Luis M. Laosa, Educational Testing Service, Princeton, New Jersey 08541, U.S.A.

The Editor forwards each manuscript to at least two, very often more, reviewers. Regardless of the category to which they belong, scientific and scholarly excellence are the main criteria used in reviewing and evaluating manuscripts.

Manuscripts are published in the same language in which they are submitted. Galley proofs along with the manuscript copyedited for printing are sent to the principal author so that the authors can proof the galleys. A policy of this journal prohibits authors from submitting manuscripts that are published elsewhere and from submitting the same manuscript for concurrent consideration to two or more publications.

Book reviews. Solicited book reviews and book-review essays are published in the Book Review Section. Book Review Editors: Rubén Ardila, Book Review Editor for Spanish-Language Books, Universidad Nacional de Colombia, Bogotá, Colombia. Jon D. Swartz, Book Review Editor for English-Language Books, Southwestern University, Georgetown, Texas, U.S.A.

Subscriptions. The *Revista Interamericana de Psicología/Interamerican Journal of Psychology* is mailed to all members of the Sociedad Interamericana de Psicología/Interamerican Society of Psychology (SIP/ISP). All persons whose professional, scholarly, or scientific work involves an interest related to any aspect of psychological research or practice are eligible for membership. Address correspondence concerning subscriptions to either of the following: SIP/ISP Individual Subscriptions, Departamento de Psicología, Universidad de Puerto Rico, Apartado 23174, Estación UPR, Río

Piedras, Puerto Rico 00931-3174. SIP/ISP Institutional Subscriptions, 124 Professional Center Parkway, San Rafael, California 94903, U.S.A. Back issues are also available from these addresses.

A *Revista Interamericana de Psicología/Interamerican Journal of Psychology*, publicação oficial da Sociedad Interamericana de Psicología/Interamerican Society of Psychology, é publicada semestralmente. A *Revista* tem como finalidade promover o intercâmbio de informação científica e profissional entre os psicólogos do hemisfério ocidental (Américas do Norte, Central e do Sul e as Caraíbas). A *Revista* tem também como objetivo facilitar a comunicação, promover maior entendimento e patrocinar a cooperação entre os cientistas do comportamento, assim como ser de interesse a todos os psicólogos do hemisfério ocidental. Com estes objetivos, a *Revista* aceita manuscritos em todas as áreas da psicologia em geral e disciplinas correlatas. Os manuscritos são aceitos em inglês, espanhol ou português.

Os manuscritos aceitos pertencem a várias categorias: relatórios de dados originais, artigos teóricos, ensaios críticos, e trabalhos metodológicos. Ocasionalmente a *Revista* publica artigos que tratam de questões políticas relevantes à Psicologia como ciência e/ou profissão. Embora a *Revista* se dedique mais à pesquisa básica e à teoria, publica também artigos mais diretamente ligados a problemas da tecnologia e da aplicação. Pesquisas que enfocam os aspectos culturais do comportamento e desenvolvimento humanos são especialmente bem recebidas, porém os manuscritos aceitos não se limitam a esses aspectos.

Endereço do Editor: Luis M. Laosa, Educational Testing Service, Princeton, New Jersey 08541, U.S.A.

O Editor manda o manuscrito para dois ou mais consultores editoriais para avaliação crítica. O critério principal na avaliação crítica dos manuscritos é a qualidade científica e acadêmica dos mesmos, independentemente da categoría a que pertencem.

Os manuscritos são publicados na língua original em que são submetidos. As provas de galé, juntamente com o manuscrito editado para impressão, são enviadas ao autor principal a fim de que o mesmo possa fazer a revisão das provas. A política editorial da *Revista* determina que os autores não podem submeter manuscritos publicados anteriormente e que não submetam simultaneamente o mesmo manuscrito a duas ou mais publicações.

Crítica de livros. Críticas de livros e ensaios críticos sobre livros, solicitados, são publicados na Seção de Crítica de Livros. Editores de Crítica de Livros: Rubén Ardila, Editor de Crítica de Livros em Espanhol, Universidad Nacional de Colombia, Bogotá, Colombia. Jon D. Swartz, Editor de Crítica de Livros em Inglês, Southwestern University, Georgetown, Texas, U.S.A.

Assinaturas. A *Revista Interamericana de Psicología/Interamerican Journal of Psychology* é enviada a todos os membros da Sociedad Interamericana de Psicología/Interamerican Society of Psychology (SIP/ISP). Todas as pessoas cujo trabalho profissional, acadêmico ou científico se relacione de qualquer modo à pesquisa ou trabalhos práticos em Psicologia, podem se candidatar como membro da Sociedade. Envie toda correspondência relativa às assinaturas aos seguintes endereços: SIP/ISP Assinaturas Para Indivíduos, Departamento de Psicología, Universidad de Puerto Rico. Apartado 23174, Estación UPR, Río Piedras, Puerto Rico 00931-3174. SIP/ISP

Assinaturas Para Instituições: 124 Professional Center Parkway, San Rafael, California 94903, U.S.A. Cópias de volumes anteriores da *Revista* podem ser obtidas nesses mesmos endereços.

Revista Interamericana de Psicología/
Interamerican Journal of Psychology

1987 Vol. 21, Nos. 1 & 2

DIRECTOR/EDITOR

Luis M. Laosa

Educational Testing Service, Princeton, New Jersey, U.S.A.

EDITORES DE RESEÑAS DE LIBROS
BOOK REVIEW EDITORS

Rubén Ardila
Universidad Nacional de Colombia, Bogotá, Colombia
Jon D. Swartz
Southwestern University, Georgetown, Texas, U.S.A.

PERSONAL EDITORIAL/EDITORIAL STAFF

Iva Barros, *Corrector de Pruebas/Copyeditor–portugués/Portuguese*
Josephine Ondetti, *Corrector de Pruebas/Copyeditor–español/Spanish*
Elsa Rosenthal, *Corrector de Pruebas/Copyeditor–inglés/English*
Robin Roth, *Secretaria/Secretary*

GERENCIA DE CIRCULACIÓN/CIRCULATION MANAGEMENT

Barbara VanOss Marín
San Rafael, California, U.S.A.

continúa/continues

JUNTA DE CONSULTORES EDITORIALES/ BOARD OF CONSULTING EDITORS

Reynaldo Alarcón N.
Instituto Nacional de Investigación y Desarrollo de la Educación, Lima, Perú

Arrigo L. Angelini
Universidade de São Paulo, São Paulo, Brasil

Stephen A. Appelbaum
Prairie Village, Kansas, U.S.A.

Rubén Ardila
Universidad Nacional de Colombia, Bogotá, Colombia

José J. Bauermeister
Universidad de Puerto Rico, Rio Piedras, Puerto Rico

Angela M. B. Biaggio
Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Amalio Blanco
Universidad Autónoma de Madrid, España

Luis Castro Bonilla
Universidad Nacional Autónoma de México, México, D.F., México

Hunter M. Breland
Educational Testing Service, Princeton, New Jersey, U.S.A.

Alvin G. Burstein
University of Tennessee, Knoxville, Tennessee, U.S.A.

Raymond Buriel
Pomona College, Claremont, California, U.S.A.

Rogelio Díaz-Guerrero
Universidad Nacional Autónoma de México, México, D.F., México

Francis J. Di Vesta
Pennsylvania State University, University Park, Pennsylvania, U.S.A.

Dorothy H. Eichorn
University of California, Berkeley, California, U.S.A.

Gustavo Fernández
Universidad Nacional Autónoma de México, Mérida, Yucatán, México

Gordon E. Finley
Florida International University, Miami, Florida, U.S.A.

contrada/continues

Otto E. Gilbert A.
Universidad del Valle de Guatemala, Guatemala, Guatemala

Ronald W. Henderson
University of California, Santa Cruz, California, U.S.A.

Wayne H. Holtzman, Sr.
University of Texas at Austin, Austin, Texas, U.S.A.

Mauricio Knobel
Universidade Estadual de Campinas, Campinas, Brasil

Federico R. León
Instituto de Psicología Aplicada al Desarrollo Económico y Social, Lima, Perú

Florente López R.
Universidad Nacional Autónoma de México, México, D.F., México

Robert B. Malmo
McGill University, Montreal, Quebec, Canada

Gerardo Marín
University of San Francisco, San Francisco, California, U.S.A.

Juracy C. Marques
Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Maritza Montero
Universidad Central de Venezuela, Caracas, Venezuela

Luiz Natalicio
University of Texas at El Paso, El Paso, Texas, U.S.A.

João Batista A. e Oliveira
Universidade de Brasília, Brasília, D.F., Brasil

Esteban L. Olmedo
California School of Professional Psychology, Los Angeles, California, U.S.A.

Juan Pascual-Leone
York University, Ontario, Canada

Harold B. Pepinsky
Ohio State University, Columbus, Ohio, U.S.A.

Beeman N. Phillips
University of Texas at Austin, Austin, Texas, U.S.A.

Karl H. Pribram
Stanford University, Stanford, California, U.S.A.

continúa/continues

CONSULTORES EDITORIALES/CONSULTING EDITORS

Manuel Ramirez III
University of Texas at Austin, Austin, Texas, U.S.A.

Emilio Ribes
Universidad Nacional Autónoma de México, Iztacala, México

Aroldo Rodrigues
Universidade Gama Filho, Rio de Janeiro, Brasil

José Miguel Salazar
Universidad Central de Venezuela, Caracas, Venezuela

Victor D. Sanua
St. John's University, Jamaica, New York, U.S.A.

Nelson Hernán Serrano J.
Quito, Ecuador

Virginia Staudt Sexton
St. John's University, Jamaica, New York, U.S.A.

Janet T. Spence
University of Texas at Austin, Austin, Texas, U.S.A.

Charles D. Spielberger
University of South Florida, Tampa, Florida, U.S.A.

Harold W. Stevenson
University of Michigan, Ann Arbor, Michigan, U.S.A.

Peter Suedfeld
University of British Columbia, Vancouver, British Columbia, Canada

Jon D. Swartz
Southwestern University, Georgetown, Texas, U.S.A.

June Louin Tapp
University of Minnesota, Minneapolis, Minnesota, U.S.A.

João Cláudio Todorov
Universidade de Brasília, Brasília, D.F., Brasil

Harry C. Triandis
University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A.

Richard R. Valencia
University of California, Santa Cruz, California, U.S.A.

Jacobo A. Varela
Montevideo, Uruguay

Daniel A. Wagner
University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.

Revista Interamericana de Psicología/
Interamerican Journal of Psychology

1987 Vol. 21, Nos. 1 & 2

Contenido/Contents

continúa/continues

PSICOLOGÍA SOCIAL Y AMBIENTAL/
SOCIAL AND ENVIRONMENTAL PSYCHOLOGY

LIBROS/BOOKS

CONSULTORES AD HOC/AD HOC REVIEWERS

Además de pedir a los miembros de la Junta de Consultores Editoriales que revisen los manuscritos presentados, el Director pide a otros psicólogos que participen, como consultores editoriales ad hoc, en el proceso de evaluar manuscritos. De esta forma, ha sido posible conseguir una base aún más amplia en lo que respecta al proceso de evaluación de manuscritos. A continuación se listan los individuos que han participado como consultores ad hoc para los manuscritos presentados a este volumen de la *Revista*.

In addition to asking members of the Board of Consulting Editors to review submitted manuscripts, the Editor asks other psychologists to participate as ad hoc reviewers in the review process. In this way, we have been able to achieve an even broader review base. The following individuals served as ad hoc reviewers for manuscripts submitted to this issue of the *Journal*.

Além de pedir que membros do Corpo de Consultores Editoriais façam as críticas dos manuscritos submetidos, o Editor solicita também que outros psicólogos participem no processo de crítica dos manuscritos, como consultores editoriais "ad hoc". Deste modo, tem sido possível obter-se uma base mais ampla para as críticas. Indicamos a seguir os nomes das pessoas que participaram como consultores editoriais "ad hoc" na crítica dos manuscritos submetidos para este volume da *Revista*.

Hortensia Amaro
Boston University School of Medicine, Boston, Massachusetts, U.S.A.

William H. Angoff
Educational Testing Service, Princeton, New Jersey, U.S.A.

Richard D. Ashmore
Rutgers University, New Brunswick, New Jersey, U.S.A.

Andrés Barona
Texas A & M University, College Station, Texas, U.S.A.

John R. Bergan
University of Arizona, Tucson, Arizona, U.S.A

Martha E. Bernal
Arizona State University, Tempe, Arizona, U.S.A.

Lillian Comas-Díaz
George Washington University, Washington, DC, U.S.A.

continúa/continues

Edward A. DeAvila
DeAvila, Duncan and Associates, San Rafael, California, U.S.A.

Rolando Díaz-Loving
Universidad Nacional Autónoma de México, México, DF. México

Richard P. Durán
University of California, Santa Barbara, California, U.S.A.

Toni L. Falbo
University of Texas at Austin, Austin, Texas, U.S.A.

Luis González de Alba
México, DF. México

Donald R. Goodenough
Rutgers University, New Brunswick, New Jersey, U.S.A.

Rafael Hernández Nieto
Universidad de los Andes, Mérida, Venezuela

Kenneth A. Holroyd
Ohio University, Athens, Ohio, U.S.A.

Spencer Kagan
University of California, Riverside, California, U.S.A.

Kathryn Kelley
State University of New York at Albany, Albany, New York, U.S.A.

Francisco J. Labrador
Universidad Complutense, Madrid, España

Herbert M. Lefcourt
University of Waterloo, Waterloo, Ontario, Canada

Barbara VanOss Marín
University of San Francisco, San Francisco, California, U.S.A.

Herbert W. Marsh
University of Sydney, Sydney, Australia

Jane R. Mercer
University of California, Riverside, California, U.S.A.

Susan Pick de Weiss
Universidad Nacional Autónoma de México, México, DF, México

Jesús Martín Ramírez
Universidad de Sevilla, Sevilla, España

continúa/continues

Ileana Recagno-Puente
Universidad Central de Venezuela, Caracas, Venezuela

Miguel Salas Sánchez
Universidad de los Andes, Bogotá, Colombia

Euclides Sánchez
Universidad Central de Venezuela, Caracas, Venezuela

David Shapiro
University of California, Los Angeles, California, U.S.A.

Milton N. Silva
Columbia Hospital, Milwaukee, Wisconsin, U.S.A.

Elaine Soto
New York, New York, U.S.A.

Deborah J. Stipek
University of California, Los Angeles, California, U.S.A.

Lawrence J. Stricker
Educational Testing Service, Princeton, New Jersey, U.S.A.

Martha L. Stocking
Educational Testing Service, Princeton, New Jersey, U.S.A.

Richard S. Surwit
Duke University Medical Center, Durham, North Carolina, U.S.A.

José A. Vigo
Brooklyn, New York, U.S.A.

Merlin C. Wittrock
University of California, Los Angeles, California, U.S.A.

Larry D. Young
Wake Forest University, Winston-Salem, North Carolina, U.S.A.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp. 1-24

Cross-cultural Measurement

FRITZ DRASGOW

and

CHARLES L. HULIN

University of Illinois at Urbana–Champaign
U.S.A.

This article discusses psychometric issues that are relevant to translations of scales. Three methods for evaluating the quality of a translation are critically reviewed. Then a translation of the Job Descriptive Index (JDI; Smith, Kendall, & Hulin, 1969) into Spanish is evaluated by one of these methods. Three groups, all employed by the same international retail marketing organization, responded to either the original English version or the translated Spanish version of the JDI. The three samples consisted of 169 bilingual U.S. Hispanics employed in New York City and Miami, 486 U.S. mainstream individuals (presumably monolingual), and 346 generally Spanish–monolingual Mexicans employed in Mexico City. Item response theory was used to assess the equivalence of measurement provided by the JDI across the three groups. In general, there were relatively few differences (11 of 68 items) for the U.S. mainstream–U.S. Hispanic comparisons. Comparisons of U.S. mainstream with Mexicans revealed 21 significant differences and there were 9 significant differences for the U.S. Hispanic – Mexican comparisons. These results differ from an earlier study (Hulin, Drasgow, & Komocar, 1982) that found only three significant differences for the same translation when responses by bilinguals were analyzed. Thus, evaluations of translations on the basis of responses by bilingual subjects are methodologically suspect. Finally, patterns of significant differences between items provide evidence for different sources (linguistic versus cultural) of measurement nonequivalence.

When can we meaningfully compare scale scores from one cultural group speaking one language to scores from another cultural group that may speak a second language? The

The authors wish to thank Frank J. Smith for his help and cooperation in obtaining the data sets analyzed in this article. He provided the initial contacts with local store managers and generally enabled us to collect data with a minimum of problems.

Copyright © 1987 by Fritz Drasgow and Charles L. Hulin. All rights reserved.

first and most obvious requirement is that the trait or attitude assessed by the scale must have some meaning in both cultures. For example, we might not expect an affective response toward one's pay to have much meaning for members of an isolated commune in a socialistic society; consequently, we should not compare their mean score on a paper-and-pencil measure of pay satisfaction (assuming that they would be willing to respond to such an instrument) to the mean score of workers in a private sector organization in a capitalistic society.

If a trait or attitude exists and has meaning for both cultural groups, it seems reasonable to ask whether the scale has a similar pattern of relations with other variables. More formally, does the trait or attitude have equivalent relations with external variables across cultures?

Before examining relations with external variables, however, it is necessary to determine whether the scale provides equivalent measurement across cultures. *Measurement equivalence* is obtained when individuals with equal standings on the trait measured by the scale, but sampled from different subpopulations, have the same expected observed score. Note that there is no requirement for equal means or variances for either the observed scores or the latent trait. Instead, measurement equivalence requires an identical relation between observed scores and the latent trait.

The reason for examining measurement equivalence prior to relational equivalence can be seen in the following example. The Job Descriptive Index (JDI; Smith, Kendall, & Hulin, 1969) includes a pay satisfaction scale that was developed for the mainstream, English-speaking workforce in the United States. Suppose an incorrect, badly garbled translation was administered to a Spanish-speaking sample. The relations of the scores of these individuals with absenteeism, turnover, and job search behaviors cannot be expected to be similar to the relations observed for English-speaking samples in the United States because the translation was faulty. Thus, a scale must provide equivalent measurement across cultures in order for the question of whether it has a similar pattern of relations with other variables to be meaningful.

Suppose, on the other hand, that we initially conducted a study of measurement equivalence and found that the scale was well translated and provided equivalent measurement across cultures and languages. If different relations of scale scores with external variables were subsequently found, they must reflect true differences between cultures in terms of the relations between the trait and external variables; the different relations could not be attributed to measurement nonequivalence.

Therefore, a scale should be examined using the two-stage procedure described by Drasgow (1984). In the first stage of this procedure, measurement equivalence should be examined. Items that exhibit divergent measurement properties, called *biased items*, can be removed from the scale. Moreover, we should determine whether individuals with equal standings on the characteristic measured by the scale have equal expected total scores (i.e., equal true scores), regardless of their group membership. In the second stage, the relations of the scale with other important variables should be examined across relevant groups. Note that nonequivalent relations would not be a result of biased items because the first stage would remove such items.

Equivalence of measurement is important for cross-cultural researchers and for complex organizations that train, evaluate, and assimilate employees from culturally heterogeneous populations who may speak different languages. Assessments that lack measurement equivalence across diverse subpopulations can result in flawed research and erroneous decisions that are costly to individuals, organizations, and society. Organizational interventions, if used where they are not needed, are costly; they detract from organizational developments that might be applied to other problems more effectively.

Individuals sampled from different cultural groups *do* respond differently to items measuring attitudes. Hulin (1987) summarized the results of several studies examining the measurement equivalence of the JDI across languages (English, French, Spanish, Hebrew, and Filipino) and countries (United States, Mexico, Canada, Philippines, and Israel). Although translation effects were found in comparisons across language within culture (e.g., bilingual Canadian Francophones responding to the original English version of the JDI versus Canadian Francophones responding to a French translation of the JDI), greater differences across culture within language were found (e.g., Canadian Anglophones versus U.S. mainstream subjects).

In the next section of this paper we review three approaches to examining scales across groups. We shall assess the extent to which these methods can both (a) detect measurement nonequivalence caused by biased items; and (b) identify nonequivalent relations with external variables. We then present an example that shows how to assess measurement equivalence. Data from U.S. mainstream, U.S. Hispanic, and Mexican subjects are used in this section. Finally, some conclusions are drawn.

METHODS FOR EXAMINING MEASUREMENT AND RELATIONAL EQUIVALENCE

In this section we critically review three methods for their powers to detect measurement and relational nonequivalence. The methods include tests of differential validity, differential prediction, and item bias.

Differential Validity

A differential validity study investigates the hypothesis that the correlation ρ_A between a scale X and another measure Y for culture A is identical to the correlation ρ_B for culture B. The alternative hypothesis is that the two correlations are not equal. *Differential validity* is said to occur when the null hypothesis of no difference is rejected.

In cross-cultural research, it certainly seems plausible that differential validity will occur when a scale includes a sizable number of biased items. Does this conclusion

necessarily follow from the premise? Surprisingly, it does not. The correlation coefficients ρ_A and ρ_B between scale X and variable Y may be almost identical even when a substantial number of scale items contain material available or familiar to only one culture (Drasgow, 1982; Drasgow & Kang, 1984). Therefore, a differential validity analysis should not be used to study measurement equivalence.

The effects of culturally biased items can be illustrated by an example presented by Drasgow (1982). Assume a scale of 100 items is developed and standardized using individuals from culture A. We would expect all 100 items to be culturally appropriate for this subpopulation. Suppose culture B overlaps with culture A to the extent that 75 items are appropriate for culture B; suppose further that culture B examinees respond randomly to the remaining 25 items, which measure material unique to culture A. If we assume that the scale mean, standard deviation, reliability, and correlation with variable Y for culture A are 60, 12, .92, and .399, and make several technical assumptions, it can be shown that the mean score in culture B is only 50, which is almost one standard deviation lower than the mean for culture A. Nonetheless, the correlation of the scale with variable Y for culture B is $\rho_B = .385$, a difference of only .014 from the culture A correlation coefficient.

The difference between the correlation coefficients for the two hypothetical subpopulations in this example is extremely small in a statistical sense. Suppose we randomly sample N_A and N_B individuals from cultures A and B, respectively, compute sample correlation coefficients, and test the hypothesis that ρ_A (= .399) equals ρ_B (= .385) at the $\alpha = .05$ level. If $N_A = 100$ and $N_B = 50$, the power of the usual significance test for the equality of correlations is only .051, which is only .001 greater than the alpha level. Doubling these sample sizes so that $N_A = 200$ and $N_B = 100$ increases the power to .052. The power of this test is still only .273 for the unreasonably large sample sizes of $N_A = 20{,}000$ and $N_B = 10{,}000$.

Note that the practical implication of the use of the differential validity analysis with ordinary sample sizes would be to conclude that no bias exists, regardless of whether the scale does or does not include biased items. In the above example, we might incorrectly accept the nearly one standard deviation difference between scores in the two cultures as meaningful. In the example of comparing the job satisfaction of workers in a commune and workers in a private sector organization, workers in the socialistic commune might truly be more satisfied with their work than workers in the capitalistic organization; however, the differential validity analysis might erroneously lead us to conclude that workers in the capitalistic culture were more satisfied because the scale included a number of items that were irrelevant for members of the socialistic commune.

The differential validity approach fares little better for the study of relational equivalence. If the variances of the scale differ across groups, then differences in correlations may reflect little more than range restriction.

Differential Prediction

In a differential prediction study, the linear regression $\hat{Y} = a_A + b_A X$ of a measure Y on scale X in culture A is compared to the regression $\hat{Y} = a_B + b_B X$ in culture B. Here we might expect measurement nonequivalence of X across cultures to cause some difference between $[a_A, b_A]$ and $[a_B, b_B]$; *differential prediction* is said to exist when we reject the hypothesis of equal regressions.

Using the same hypothetical test and making the same assumptions described above, it can be shown that differential prediction analyses frequently have far more statistical power than differential validity analyses (Drasgow, 1982; Drasgow & Kang, 1984). Unfortunately, Drasgow and Kang found that even the differential prediction method requires fairly large sample sizes to have adequate power; samples of at least several hundred seem to be necessary. Thus, the sample sizes necessary for such a study to be technically feasible are much larger than previously believed and larger than frequently encountered.

The superior power of the differential prediction analysis for detecting measurement nonequivalence can be illustrated with the hypothetical 100-item scale described previously. When the sample sizes of the culture A and culture B are $N_A = 100$ and $N_B = 50$ and $\alpha = .05$, the power of the differential prediction significance test is .42. If $N_A = 200$ and $N_B = 100$, the power is .72, and the power is approximately 1.00 when $N_A = 20{,}000$ and $N_B = 10{,}000$. Unfortunately, sample sizes larger than $N_A = 200$ and $N_B = 100$ are probably required to achieve the powers listed here because our assumption that members of culture B respond randomly to the items unique to culture A is probably too strong.[1]

Although it is more powerful than the differential validity analysis, the differential prediction analysis still has a serious limitation. To illustrate this point, suppose that we find cross-cultural differences in regressions. From the differential prediction analysis alone, it is impossible to know whether the differences are due to measurement nonequivalence or relational nonequivalence; these two effects are confounded.

Another problem with the differential prediction analysis is that the criterion measure Y may contain elements that are unique to one culture. In some settings, these elements may be integral parts of the criterion; when this is true it illustrates the difficulties of cross-cultural research. In this situation, a test or scale that provided equivalent measurement by measuring only common concepts may be inadequate for predicting other measures with large unique components. Two individuals with the same score on such a scale, but sampled from the two different cultures, would not be expected to have the same score on Y.

IRT Item Bias Analyses

The differential validity and differential prediction analyses do not *directly* assess measurement equivalence; they presume that nonequivalent measurement will lead to

[1]Random responses imply horizontal item characteristic curves and no relation between item responses and the latent trait. A more plausible model would incorporate monotonically increasing item characteristic curves for culture B that were displaced to the right of culture A's item characteristic curves and were flatter.

differential X–Y relations, which in turn can be detected. As noted above, this presumption is essentially false for the differential validity analysis and is difficult to evaluate for the differential prediction analysis.

The IRT item bias analysis, in contrast, directly tests the equivalence of measurement provided by individual items. In particular, this analysis tests whether examinees with equal standings on the latent trait θ assessed by the scale, but sampled from different groups, have the same chance of responding positively to an item. If this analysis is performed first and items and scales are found to provide equivalent measurement, then the results of the differential prediction analysis can be used directly to interpret relational equivalence.

A fundamental notion in an item bias analysis is the item characteristic curve (ICC), which is a plot of the conditional probability of a correct or positive response as a function of the latent trait θ. Illustrations of ICCs showing item bias across subpopulations can be found in Hulin, Drasgow, and Parsons (1983, p. 176).

There are two assumptions made by most item response theories. First, it is assumed that all items measure a single latent trait and therefore the item pool is unidimensional. Methods for checking this assumption using binary item response data are described in Chapter 8 of Hulin et al. (1983). The second assumption is that the probabilities of correct responses at different ability levels can be fitted adequately by some IRT model. The three-parameter logistic model, originally developed by Birnbaum (1968) and extensively studied by Lord (1980), is frequently fitted to multiple-choice test items. A method for checking the fit of such a model to actual item response data is illustrated by Figure 2.4.3 of Hulin et al. (1983, p. 30).

Lord's (1980) approach to an IRT item bias study consists of the following steps. (See Thissen, Steinberg, & Gerrard, 1986, for an alternative IRT approach.) First, item parameters should be estimated separately from sample data obtained from the two subpopulations. A computer program such as LOGIST (Wood, Wingersky, & Lord, 1976) or BILOG (Mislevy & Bock, 1985) can be used for this purpose. Then the assumptions concerning unidimensionality and the fit of the model to the data should be checked. If the assumptions are reasonably satisfied, then it is necessary to "link" the latent trait metrics for the two samples. Metrics for the two samples must be linked because the θs (or θ estimates) for each of the two groups are expressed as z scores (means of zero, standard deviations of one), and then item parameter estimates are expressed in relation to the latent trait metric. Note that real differences between the two groups can be hidden by scaling θ metrics as z scores within each group. Linking rescales the θ and item parameter estimates of the two groups so that they are expressed in the same scale. Hulin et al. (1983, p. 202), Stocking and Lord (1983), and Segall (1983) all describe and illustrate methods for linking metrics. Finally, an item bias statistic can be computed for each item by the formula presented by Lord (1980, Chapter 14). Each item bias value provides a quantitative measure of the difference between estimated ICCs for the two samples; the corresponding significance test is used to determine whether the observed difference between sample ICCs is due to

sampling variance. Items with large item bias statistics can then be identified as biased; they fail to provide equivalent measurement across the two subpopulations.

Lord (1977) conducted one of the earliest item bias studies using IRT. He compared item responses of U.S. Whites and Blacks to the April, 1975 Scholastic Aptitude Test, Verbal Section. Lord found that 38 of the 85 items provided nonequivalent measurement (when the alpha level of the significance test was .05). The very large sample sizes used by Lord ($N = 2,250$ for each group) certainly contributed to the number of items found to be biased. When sample sizes are this large, trivial differences between ICCs are often statistically significant. Unfortunately Lord did not indicate how the large number of biased items would affect test scores across the two groups; the question of total score or scale score bias was not addressed.

The IRT item bias studies to date have generally found substantial proportions of items with relatively large bias statistics. These statistics must be viewed with some caution, however. Minor violations of unidimensionality or of the hypothesized parametric form of the ICCs may cause item bias statistics to be inflated. Also, Lord's (1980) expressions for the standard errors of item parameter estimates obtained with the LOGIST computer program substantially underestimate the actual standard errors (McLaughlin & Drasgow, 1987). In contrast, Drasgow (1988) and Lim (1987) have found much more accurate standard errors when item parameters were estimated by a program (e.g., BILOG) that used the method of marginal maximum likelihood.

Finally, one other aspect of IRT item bias studies should be noted. The *test characteristic curve* (TCC), which is the sum over items of the ICCs, shows the relation between expected observed scores (i.e., true scores) and the trait measured by the scale. TCCs present evidence about measurement bias at the total scale score level, which is very important because total scores are normally used to draw inferences about individuals and groups. Appropriately constructed TCCs for groups from two cultures show the extent of the unique components of a scale: If a scale has many items with content specific to culture A, then the TCC for culture A would lie above the TCC for culture B, and the difference between the two TCCs at a given θ shows how much the scale favors culture A.

One must, however, be careful when comparing TCCs: Methods of linking θ metrics may force TCCs to be very similar across groups regardless of the extent of measurement nonequivalence at the item level. To circumvent this problem, we temporarily discard items found to be biased, and then relink θ metrics using only items found to be unbiased. This process improves the linking because it removes items that seem to be measuring differently in the two groups. After we have relinked metrics on the basis of items found to be unbiased, all item parameter estimates can be expressed in terms of this new metric, and all items can again be tested for bias. This procedure can be iterated until the same items are found to be biased on two successive iterations. Thus, the final linking is based only on items found to be unbiased. In this way consistent bias on a subset of items can be detected and the TCCs for the rescaled item parameter estimates show the magnitude of the bias.

MEASUREMENT EQUIVALENCE FOR U.S. MAINSTREAM, U.S. HISPANICS, AND MEXICANS

Measurement characteristics of the JDI in three distinct groups are described in this section. The JDI is used here because this instrument was explicitly designed to assess facets of job satisfaction "in a wide range of occupational and educational groups" (Smith et al., 1969, p. 71). It is also generally regarded as the most carefully developed measure of job satisfaction in existence today (Campbell, 1970; Robinson, Athanasiou, & Head, 1968; Vroom 1964). It is clear that Smith et al. went to great lengths to ensure the equivalence of their measurement operations across various subpopulations. The JDI should exhibit measurement equivalence if this is a property that can be obtained through diligent efforts. If the JDI does not exhibit equivalence, then the vast majority of instruments currently in use by applied psychologists that were developed by less rigorous methods are unlikely to have this important property.

In the rest of this section we shall describe an empirical study of the equivalence of measurement provided by the JDI. Samples from three distinct cultures (U.S. mainstream, U.S. Hispanic, Mexican) responded either to the original English version of the JDI or to a translation of the JDI into Spanish.

Method

Subjects, administrative procedures, and measures. All subjects were employees of an international retail sales firm. The subjects were primarily nonmanagerial employees who worked in a variety of job types such as sales, clerical, and maintenance.

The U.S. mainstream sample consisted of 486 employees taken from a larger sample of 1,046 respondents with complete data. These subjects completed a written survey in groups of 10 to 50 individuals. The survey was administered by organizational staff on company time. The completed surveys were sealed in envelopes and mailed to university researchers. Subjects were assured of the confidentiality of their responses and of the voluntary nature of their participation in the research. Although no data concerning language skills were collected, it seems safe to assume that the vast majority of these subjects were monolingual English speakers.

The U.S. Hispanic sample consisted of 169 bilingual subjects employed in the Miami or New York City areas with complete data. These subjects were of Cuban or Puerto Rican ethnic backgrounds and were fluent in both English and Spanish. They completed both English and Spanish versions of a written attitude questionnaire on two occasions in a counterbalanced order. The second administration was one month after the first.

The Mexican sample consisted of 346 subjects with complete data employed in Mexico City. These employees were selected from the non-supervising workforce of

the same organization that provided the U.S. mainstream and U.S. Hispanic samples. The questionnaires were administered by a researcher from the University of Illinois who was accompanied by a member of the local Mexican branch of the organization. The research nature of the study was stressed, and confidentiality was assured. Other details of the administration were as similar to the U.S. samples as local circumstances and facilities would permit. Most of these subjects were monolingual Spanish-speaking employees.

The five JDI scales (Work, Supervision, Pay, Promotion Opportunities, Coworkers) were included in the questionnaire administered to each sample. The U.S. mainstream sample completed the original (English) form of the JDI. The U.S. Hispanic sample completed both the English version of the JDI as well as a translation of the JDI into Spanish. In previous analyses (Hulin, Drasgow, & Komocar, 1982) of the U.S. Hispanic data set, we found that only three of the 72 JDI items had significantly different ($\alpha = .05$) ICCs across the two languages. Consequently, all analyses described in this section used responses of the U.S. Hispanic sample to the Spanish version of the JDI. Finally, the Mexican sample completed the Spanish version of the JDI.

Analyses. Two-parameter normal ogive ICCs (Lord, 1952) were used to model the JDI data. In this model, the probability of a positive response to item i among subjects with attitude θ is

$$(1) \qquad P_i(\theta) = \int_{-\infty}^{g_i} \phi(t)\, dt,$$

where $\phi(t)$ is the height (i.e., the density) of the normal curve at t:

$$(2) \qquad \phi(t) = (2\pi)^{-1/2} \exp(-t^2/2),$$

and the upper limit of the integral is

$$(3) \qquad g_i = a_i(\theta - b_i).$$

In Equation 3, a_i is the item discriminating power (i.e., the parameter that controls the steepness of the ICC) and b_i is the item extremity (i.e., the point along the θ continuum at which subjects respond positively with probability .5). Previous research (Hulin et al., 1983, Chapter 3) has indicated that a lower asymptote parameter (often denoted as c_i) is probably not necessary for JDI items.

The two-parameter normal ogive model is a unidimensional IRT model in the sense that the attitude parameter θ is a scalar. Therefore, we analyzed each of the five JDI scales separately in order to reduce multidimensionality. IRT analyses of 9 and 18 item scales is problematical with estimation methods that use the method of "conditional" maximum likelihood (e.g., LOGIST; see Lord, 1968); consequently, we used the marginal maximum likelihood approach described by Bock and Lieberman (1970) and programmed by Drasgow (1988). This program can be used to analyze scales with as few as five items.

After obtaining item parameter estimates for all five scales for the three samples (a total of 15 scale calibrations), item parameters for each scale were linked. The Stocking–Lord (1983) method (with the 99 centile points of the normal distribution in place of attitude estimates) was used to link metrics. Then a χ^2 item bias statistic (see

Lord, 1980, Chapter 14) was computed for each item to test simultaneously the hypothesis that item discrimination a_i and item extremity b_i are identical across two groups. Note that if these parameters are identical in two subpopulations, then the ICCs must also be identical.

The Stocking–Lord linking method was then reapplied to each scale using only the items that were previously found to be unbiased. This yielded new linking constants, which were then applied to all items on the scale (including the ones previously identified as biased), and the χ^2 item bias statistic was recomputed for each item. This procedure continued until the same items were identified as biased in two successive iterations.

Results

Tables 1 through 5 present the results of the item bias analyses. The results for the Work Satisfaction scale are shown in Table 1. The first column under the heading "χ^2" shows the significant χ^2 statistics in the U.S. mainstream–U.S. Hispanic comparisons. Note that only three of the 17 items provided nonequivalent measurement across the two subpopulations. The second column in Table 1 shows that four items produced nonequivalent measurement in the U.S. Hispanic and Mexican samples. Finally, the third column gives the results for the U.S. mainstream–Mexican comparisons. Here seven items were found to provide nonequivalent measurement. ICCs for an item (*creative*) found to be unbiased in all comparisons are shown in Figure 1. Note that the probabilities of a positive response to this item are very similar at all θ values for all three groups.

One of the items (*challenging*) on the Work Satisfaction scale was previously identified as mistranslated (Hulin et al., 1982, p. 823) and, consequently, was expected to yield nonequivalent measurement. ICCs for this item are presented in Figure 2. This item is quite discriminating for the U.S. mainstream sample, is relatively undiscriminating for the U.S. Hispanic sample, and provides almost no information about work satisfaction for the Mexican sample. Clearly, the item has very different measurement properties in the three groups.

Table 1
Item Parameter Estimates for the Work Satisfaction Scale

Item	χ^2			U.S. Mainstream		U.S. Hispanic		Mexican	
	1	2	3	a_i	b_i	a_i	b_i	a_i	b_i
Fascinating		19	49	1.13	−.02	1.05	.23	.80	1.01
				(.14)	(.08)	(.23)	(.14)	(.11)	(.12)
Routine				.66	.49	.66	.05	.94	.56
				(.10)	(.13)	(.16)	(.17)	(.14)	(.11)
Satisfying		19	27	1.84	−1.05	1.71	−1.06	1.25	−.44
				(.26)	(.07)	(.43)	(.15)	(.21)	(.13)

Item Parameter Estimates for the Work Satisfaction Scale (Cont.)

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Boring				1.38	−1.25	1.05	−1.37	1.73	−.82
				(.18)	(.09)	(.25)	(.23)	(.39)	(.14)
Good				.96	−1.80	1.59	−1.26	1.03	−1.28
				(.14)	(.17)	(.41)	(.18)	(.20)	(.23)
Creative				1.00	−.15	.76	.00	.81	−.22
				(.12)	(.08)	(.17)	(.16)	(.12)	(.15)
Respected				.99	−1.10	.66	−1.38	.77	−1.64
				(.12)	(.10)	(.16)	(.32)	(.15)	(.33)
Hot		21	27	.25	−1.30	.34	−.91	.27	.91
				(.08)	(.35)	(.12)	(.37)	(.07)	(.26)
Pleasant			18	.87	−1.36	1.11	−.93	1.05	−.53
				(.10)	(.11)	(.21)	(.15)	(.17)	(.15)
Useful				.71	−2.21	.61	−1.69	1.01	−1.71
				(.11)	(.23)	(.15)	(.33)	(.22)	(.30)
Tiresome		13		.83	−.31	.56	−.85	.54	.03
				(.11)	(.08)	(.14)	(.25)	(.09)	(.17)
Healthful				.41	.97	.63	−.07	.82	.56
				(.09)	(.32)	(.15)	(.18)	(.12)	(.12)
Challenging	40		59	1.21	−1.04	.53	.89	.16	1.41
				(.15)	(.09)	(.14)	(.30)	(.06)	(.50)
Frustrating	31		40	.45	.22	1.03	−.76	.82	−1.03
				(.09)	(.16)	(.22)	(.16)	(.14)	(.23)
Simple				.41	−1.07	.13	−.06	.42	−.82
				(.08)	(.20)	(.11)	(.77)	(.08)	(.30)
Endless				.34	.93	.31	−.02	.26	.10
				(.08)	(.30)	(.12)	(.32)	(.07)	(.31)
Gives sense of accomp.	16		21	1.81	−1.08	.77	−1.07	.60	−2.05
				(.25)	(.07)	(.16)	(.22)	(.13)	(.47)

Note: All parameters linked to a common metric by iterative Stocking–Lord linking. The first χ^2 test presents the significant ($\alpha = .005$) χ^2 statistics for the U.S. mainstream–U.S. Hispanic comparisons, the second gives the U.S. Hispanic–Mexican comparisons, and the third gives the U.S. mainstream–Mexican comparisons. The item "on your feet" was excluded because its item parameters could not be estimated accurately.

Table 2

Item Parameter Estimates for the Satisfaction With Supervision Scale

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Asks my advice	22		22	.71	−.69	.66	.27	.42	.16
				(.09)	(.11)	(.15)	(.17)	(.08)	(.18)

Item Parameter Estimates for the Satisfaction With Supervision Scale (Cont.)

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Hard to please				1.40	−.97	1.19	−.66	1.32	−.64
				(.21)	(.12)	(.22)	(.14)	(.16)	(.10)
Impolite				1.32	−1.29	1.51	−.86	.92	−1.45
				(.17)	(.11)	(.31)	(.13)	(.12)	(.14)
Praises good work				1.17	−.56	1.06	−.71	1.20	−.91
				(.13)	(.08)	(.21)	(.16)	(.16)	(.11)
Tactful				1.16	−.48	1.26	−.43	1.75	−.61
				(.13)	(.08)	(.23)	(.13)	(.27)	(.09)
Influential		16	39	.93	−.24	.63	−.63	.15	1.37
				(.12)	(.08)	(.15)	(.21)	(.05)	(.55)
Up-to-date	28	30		1.46	−.89	.54	−.66	1.02	−1.25
				(.17)	(.08)	(.13)	(.24)	(.13)	(.12)
Doesn't supervise enough		12		.78	−.74	.42	−.82	.76	−.73
				(.10)	(.12)	(.11)	(.30)	(.10)	(.12)
Quick tempered				.89	−.92	1.42	−.61	1.16	−.67
				(.12)	(.12)	(.29)	(.14)	(.16)	(.10)
Tells me where I stand			20	.60	−.25	.82	−.58	.80	−.84
				(.09)	(.11)	(.16)	(.17)	(.11)	(.13)
Annoying				2.04	−.93	2.35	−.94	1.31	−.32
				(.32)	(.08)	(.67)	(.13)	(.17)	(.10)
Stubborn	13			1.09	−.88	1.09	.30	.92	−.64
				(.13)	(.11)	(.20)	(.13)	(.12)	(.11)
Knows job well			13	1.18	−.97	.75	−1.63	1.20	−1.48
				(.14)	(.10)	(.17)	(.29)	(.18)	(.14)
Bad				1.46	−1.41	1.67	−1.31	1.63	−1.32
				(.20)	(.11)	(.44)	(.19)	(.25)	(.11)
Intelligent				.90	−1.23	1.13	−1.18	.98	−1.15
				(.11)	(.13)	(.24)	(.19)	(.12)	(.11)
Leaves me on my own	a		a	a	a	.66	−1.26	.96	−.91
						(.13)	(.25)	(.12)	(.11)
Lazy		17		1.10	−1.76	.49	−2.26	1.34	−1.62
				(.18)	(.18)	(.14)	(.62)	(.20)	(.14)
Around when needed				.80	−.82	.96	−1.05	.99	−1.08
				(.11)	(.12)	(.21)	(.19)	(.13)	(.11)

[a]Item parameters could not be estimated for the U.S. mainstream sample.

Note: All parameters linked to a common metric by iterative Stocking–Lord linking. The first χ^2 test presents the significant ($\alpha = .005$) χ^2 statistics for the U.S. mainstream–U.S. Hispanic comparisons, the second gives the U.S. Hispanic–Mexican comparisons, and the third gives the U.S. mainstream–Mexican comparisons.

Table 3
Item Parameter Estimates for the Pay Satisfaction Scale

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Income adequate for normal expenses				1.26 (.14)	.22 (.07)	.83 (.17)	.40 (.15)	1.50 (.21)	.26 (.08)
Satisfactory profit sharing	a	b	a	a	a	b	b	.79 (.12)	1.02 (.16)
Barely live on income				1.43 (.16)	.36 (.07)	1.17 (.22)	.60 (.13)	1.10 (.15)	.49 (.10)
Bad				2.20 (.33)	.01 (.06)	2.11 (.51)	−.24 (.12)	2.05 (.34)	.12 (.07)
Income provides luxuries	a	a	a	.78 (.11)	1.59 (.15)	a	a	a	a
Insecure				.98 (.12)	−.12 (.09)	1.16 (.22)	−.07 (.13)	.74 (.23)	−.40 (.09)
Less than I deserve				1.66 (.20)	.72 (.06)	1.39 (.27)	.82 (.12)	1.58 (.23)	.66 (.09)
Highly paid				.93 (.19)	2.66 (.31)	.97 (.25)	1.83 (.30)	1.44 (.30)	1.61 (.17)
Underpaid	a	a	41	2.33 (.35)	.73 (.06)	a	a	1.99 (.32)	.13 (.08)

Note: All parameters linked to a common metric by iterative Stocking–Lord linking. The first χ^2 test presents the significant ($\alpha = .005$) χ^2 statistics for the U.S. mainstream–U.S. Hispanic comparisons, the second gives the U.S. Hispanic–Mexican comparisons, and the third gives the U.S. mainstream–Mexican comparisons.

[a] Item parameters could not be estimated in this sample.

[b] Item was not included on survey.

Table 4
Item Parameter Estimates for the Satisfaction With Promotion Opportunities Scale

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Good opportunity for advancement				3.58 (.63)	.35 (.06)	3.35 (.94)	.16 (.11)	2.33 (.36)	.22 (.07)
Opportunity somewhat limited				1.36 (.16)	.77 (.08)	1.22 (.22)	.63 (.15)	1.55 (.20)	.64 (.08)
Promotion on ability				1.38 (.15)	.06 (.06)	1.40 (.27)	−.17 (.12)	1.02 (.14)	−.04 (.11)
Dead-end job				1.30 (.14)	−.13 (.07)	1.52 (.28)	−.06 (.12)	1.71 (.23)	.15 (.08)
Good chance for promotion				4.07 (.73)	.31 (.05)	3.49 (.95)	.29 (.11)	4.20 1.29	.36 (.06)

Table 4
Item Parameter Estimates for the Satisfaction With Promotion Opportunities Scale (Cont.)

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Unfair promotion policy			37	.81	−.34	1.19	−.03	1.42	.36
				(.10)	(.09)	(.21)	(.13)	(.18)	(.08)
Infrequent promotions			19	.96	.56	1.18	.89	1.30	.96
				(.11)	(.09)	(.24)	(.17)	(.17)	(.09)
Regular promotions		18	49	1.60	.80	1.55	.75	.17	1.97
				(.19)	(.08)	(.31)	(.14)	(.08)	(.77)
Fairly good chance for promotion			77	3.42	.18	1.88	.39	1.87	.99
				(.61)	(.05)	(.37)	(.12)	(.27)	(.08)

Note: All parameters linked to a common metric by iterative Stocking–Lord linking. The first χ^2 test presents the significant ($\alpha = .005$) χ^2 statistics for the U.S. mainstream–U.S. Hispanic comparisons, the second gives the U.S. Hispanic–Mexican comparisons, and the third gives the U.S. mainstream–Mexican comparisons.

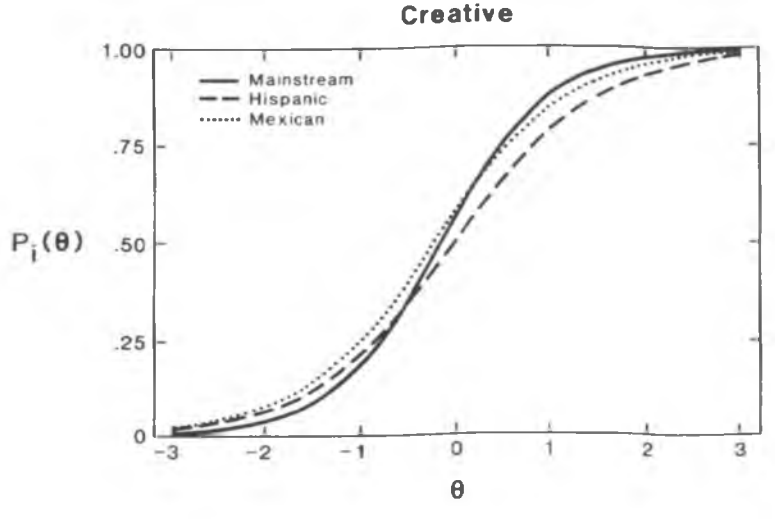

Figure 1. ICCs for an item (*creative*) from the JDI Work Satisfaction scale found to be unbiased in all comparisons: U.S. mainstream, U.S. Hispanic, and Mexican samples.

Table 5
Item Parameter Estimates for the Satisfaction With Coworkers Scale

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Stimulating				.87	−.30	.82	−.44	1.02	−.65
				(.11)	(.09)	(.15)	(.17)	(.13)	(.10)

Item Parameter Estimates for the Satisfaction With Coworkers Scale (Cont.)

Item	χ^2 1	2	3	U.S. Mainstream a_i	b_i	U.S. Hispanic a_i	b_i	Mexican a_i	b_i
Boring				1.33	−1.31	1.33	−1.04	1.20	−1.52
				(.19)	(.11)	(.27)	(.15)	(.17)	(.13)
Slow				1.18	−.68	.87	−.81	1.00	−.88
				(.14)	(.08)	(.15)	(.17)	(.13)	(.11)
Ambitious	19		39	1.01	−.24	.38	−.13	.18	−.52
				(.11)	(.08)	(.10)	(.28)	(.07)	(.37)
Stupid				1.68	−1.53	1.30	−1.21	1.02	−1.44
				(.27)	(.12)	(.28)	(.17)	(.14)	(.11)
Responsible				1.30	−1.02	1.05	−1.44	1.40	−1.20
				(.16)	(.10)	(.23)	(.22)	(.20)	(.10)
Fast			30	.91	−.05	.95	−.43	1.27	−.67
				(.11)	(.09)	(.17)	(.16)	(.17)	(.09)
Intelligent				.95	−1.18	1.16	−.83	1.46	−.95
				(.10)	(.10)	(.22)	(.15)	(.20)	(.10)
Easy to make enemies	13			1.01	−1.10	.51	−1.10	.80	−.91
				(.13)	(.11)	(.10)	(.25)	(.10)	(.12)
Talk too much	a		a	a	a	.41	−.80	.52	−.39
						(.10)	(.32)	(.09)	(.14)
Smart			24	.94	−.74	1.50	−.88	1.47	−1.17
				(.12)	(.10)	(.32)	(.14)	(.21)	(.10)
Lazy				1.24	−1.07	.75	−1.15	1.50	−1.13
				(.16)	(.10)	(.14)	(.21)	(.21)	(.09)
Unpleasant				1.20	−1.64	1.01	−1.68	1.44	−1.55
				(.18)	(.15)	(.23)	(.25)	(.22)	(.12)
No privacy	21		34	.58	−1.49	.47	−.35	.55	−.24
				(.08)	(.19)	(.11)	(.24)	(.09)	(.14)
Active				1.30	−.93	1.19	−1.08	1.50	−1.02
				(.16)	(.09)	(.24)	(.18)	(.20)	(.09)
Narrow interests	20			.84	−.50	.85	.36	.79	−.22
				(.10)	(.10)	(.16)	(.17)	(.10)	(.11)
Loyal				.89	−.68	.86	−.47	.87	−.36
				(.11)	(.10)	(.16)	(.17)	(.11)	(.11)
Hard to meet	33		19	.77	−1.74	.64	−.58	.79	−1.02
				(.10)	(.16)	(.13)	(.20)	(.11)	(.12)

Note: All parameters linked to a common metric by iterative Stocking–Lord linking. The first χ^2 test presents the significant ($\alpha = .005$) χ^2 statistics for the U.S. mainstream–U.S. Hispanic comparisons, the second gives the U.S. Hispanic–Mexican comparisons, and the third gives the U.S. mainstream–Mexican comparisons.

[a] Due to a clerical error, this item was omitted from the survey.

The pattern of significant differences in Table 1 is informative. For example, the item, *challenging*, had significant U.S. mainstream–U.S. Hispanic and U.S.

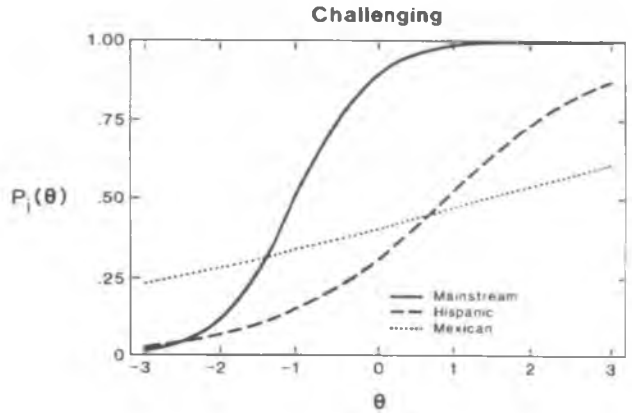

Figure 2. ICCs for an item (*challenging*) from the JDI Work Satisfaction scale found to be biased in both English–Spanish comparisons: U.S. mainstream, U.S. Hispanic, and Mexican samples.

mainstream–Mexican differences. The U.S. Hispanic–Mexican comparison, however, was not significant, a fact that supports the assertion that a translation problem created the English/Spanish differences. The items *frustrating* and *gives sense of accomplishment* had similar patterns. These items had different measurement characteristics across languages, but they provided similar measurement for the two Hispanic subpopulations. The items *fascinating*, *satisfying*, and *hot* also showed a consistency: They provided equivalent measurement across the two U.S. subpopulations but nonequivalent measurement in both U.S.–Mexican comparisons. This result suggests a cultural difference between U.S. Hispanics and Mexicans but not a translation problem (because there was no difference between ICCs estimated from the U.S. samples responding in English and Spanish). Both of these types of consistency support the notion that U.S. Hispanics share parts of both the U.S. mainstream and the Mexican cultures. Only one item (*tiresome*) was inconsistent with this hypothesis in that it had a significant U.S. Hispanic–Mexican difference but an insignficant U.S. mainstream–Mexican difference. An explanation for this anomalous finding is not readily to be found.

The results for the Satisfaction With Supervision scale are shown in Table 2. For this scale there were three significant U.S. mainstream–U.S. Hispanic differences, four significant U.S. Hispanic–Mexican differences, and only four significant U.S. mainstream–Mexican differences. Perhaps the most noteworthy aspect of Table 2 is the relatively small number of items that provided nonequivalent measurement.

Table 3 presents the results for the Pay Satisfaction scale. Only one significant χ^2 statistic is shown in this table. However, the item *satisfactory profit sharing* would have shown a significant U.S. mainstream–Mexican difference because its *a* parameter was very small for the Americans (about .1).[2] The item extremity parameter of the item *income provides luxuries* could not be accurately estimated in either the U.S. Hispanic

[2]The low *a* parameter for this item caused estimation of the location of the item extremity parameter *b* to be problematical: The ICC is essentially a horizontal line with height .5 for $-3 < \theta < 3$.

or Mexican samples. This item has a large $\hat{b}$ (= 1.59) value in the U.S. mainstream sample; because the latter tended to be more satisfied with their pay than either Hispanic group, parameters of items with large b values were more difficult to estimate in the Hispanic samples (because there were fewer subjects whose responses provided information about the parameter). Thus, it is not surprising that parameters for this item could not be estimated accurately, a fact consistent with the hypothesis that the item provides equivalent measurement across all three subpopulations. Finally, the $\hat{b}$ of the item *underpaid* for the U.S. Hispanic sample was .18, which is very similar to the Mexican $\hat{b}$ of .13. The a parameter, however, could not be estimated accurately, apparently because the sample was too small. The data do indicate that this item had a large a parameter, although we were not able to obtain a reasonable point estimate.

Table 4 contains the results for the Satisfaction With Promotion Opportunities scale. Note that there were no significant U.S. mainstream–U.S. Hispanic differences and only one significant U.S. Hispanic–Mexican difference. Nonetheless, almost half of the U.S. mainstream–Mexican comparisons were significant. Examination of Table 4 shows that the U.S.-Hispanic item parameter estimates were intermediate between the U.S. mainstream and Mexican estimates. For example, the item *unfair promotion policy* had a estimates .81, 1.19, and 1.42 and b estimates –.34, –.03, and .36 across the U.S. mainstream, U.S. Hispanic, and Mexican samples. This pattern of results suggests a cultural effect rather than translation error.

Results for the Satisfaction With Coworkers scale are shown in Table 5. There were no significant U.S. Hispanic–Mexican differences, five significant U.S. mainstream – U.S. Hispanic differences, and five significant U.S. mainstream–Mexican differences. Apparently, the U.S. Hispanics' views about interpersonal relations with peers had been relatively unaffected by their interactions with U.S. mainstreamers; they seemed to reflect the values of the Spanish culture rather than the U.S. mainstream culture.

The results shown in Tables 1 through 5 differ sharply from the results of the earlier study by Hulin et al. (1982). Hulin et al. compared the responses of the bilingual U.S. Hispanic sample when they responded in English and when they responded in Spanish. Only three items were found to be biased in this analysis, a result that led Hulin et al. to be quite optimistic about the ease of translating a scale and obtaining measurement equivalence across cultures and languages. In the present analysis, we found striking differences when the U.S. Hispanic sample was compared with the U.S. mainstream sample and when the U.S. Hispanic sample was compared with the Mexican sample. Thus, it is clear that the bilingual sample differs from either monolingual sample, and data from bilinguals responding in both languages are not adequate for evaluating a translation of a scale.

Finally, the TCCs for the three groups on the Work Satisfaction scale are shown in Figure 3. The largest difference occurs at $\theta = -0.4$. Here a U.S. mainstream subject would be expected to choose the positively keyed option on 10.6 items whereas a Mexican subject would be expected to choose the keyed response on 9.2 items. This difference corresponds to a four-point difference when standard JDI scoring is used (which assigns three points to the keyed response), or just under half of the standard deviation of the Work Satisfaction scale that was reported by Smith et al. (1969) in their

national norming of the JDI.

The differences in the TCCs displayed in Figure 3 are due to biased items such as the item *challenging* (see Figure 2). The differences are large enough that they could cause a mean difference in observed scale scores even if there were truly no mean difference in job satisfaction for Mexicans and U.S. mainstream. It should be stressed that removing biased items from the scale would eliminate the source of the problem; retaining only unbiased items would yield a scale that provides equivalent measurement and allows valid comparisons across cultural groups. Alternatively, Table 6 (which was constructed from Figure 3) could be used to transform scores for the two Hispanic groups so that their transformed scores would be directly comparable to scores of U.S. mainstreamers. For example, a score of 10 for a Mexican would be considered equivalent to a score of 11.36 for a U.S. mainstreamer. Transforming scores — rather than eliminating biased items — should be used only when the degree of bias is relatively small (although possibly statistically significant).

Table 6

Conversion Table for Transforming Scores of the Hispanic Groups to the Scale of the U.S. Mainstream Group

U.S. Hispanic		Mexican	
Raw score	**Transformed score**	**Raw score**	**Transformed score**
2	1.90	2	1.90
3	2.96	3	3.00
4	4.01	4	4.21
5	5.07	5	5.54
6	6.15	6	6.89
7	7.25	7	8.19
8	8.34	8	9.34
9	9.41	9	10.39
10	10.44	10	11.36
11	11.44	11	12.29
12	12.40	12	13.16
13	13.33	13	13.97
14	14.22	14	14.74
15	15.12	15	15.49
16	16.06	16	16.30

DISCUSSION

Although researchers have been concerned about the effects of culture on psychological measurement for many years, there have been serious problems with statistical analyses. Cross-cultural comparisons of correlations of a scale with other variables (the differential validity analysis) is unable to identify serious violations of measurement

equivalence and relational equivalence. Conclusions from regression equations (the differential prediction analysis) are limited because measurement nonequivalence is confounded with relational nonequivalence. With IRT, we can now directly evaluate measurement equivalence; once measurement equivalence is attained (by eliminating biased items or using a conversion table), we can evaluate relational equivalence through a regression analysis.

We have provided an example of how to evaluate measurement equivalence using Lord's (1980) approach. Researchers may also consider the Mantel–Haenszel procedure developed by Holland and his colleagues (e.g., Holland & Thayer, 1986) and the method illustrated by Thissen et al. (1986). All of these methods are designed to determine whether individuals with equal standings on the trait assessed by a scale, but sampled from different groups, have equal chances of responding positively to an item.

These approaches to the study of measurement equivalence provide very powerful detection of differences between groups when sample sizes are large. In fact, with samples of $N = 1{,}500$ U.S. Blacks, Whites, and Hispanics, Drasgow (1987) obtained statistically significant differences where the "effect sizes" were so small as to be of no practical consequence. The differences obtained in our comparisons of U.S. mainstreamers and Mexicans are large enough to have practical consequences. It seems likely that comparisons of groups that speak different languages and live in different countries will ordinarily exhibit substantial differences.

In this article we attempted to determine whether item bias was due to translation error or cultural difference. We did not attempt to ascertain why some items appeared to reflect cultural differences. Specifically, what are the cultural differences that cause some items to function differently across groups? Although an IRT analysis can help to identify culturally biased items, it cannot determine *why* the items are biased. Thus, an IRT item bias analysis can serve as a powerful aid by *identifying* similarities and differences across cultures, but *explaining* these similarities and differences is a task for the cross-cultural researcher.

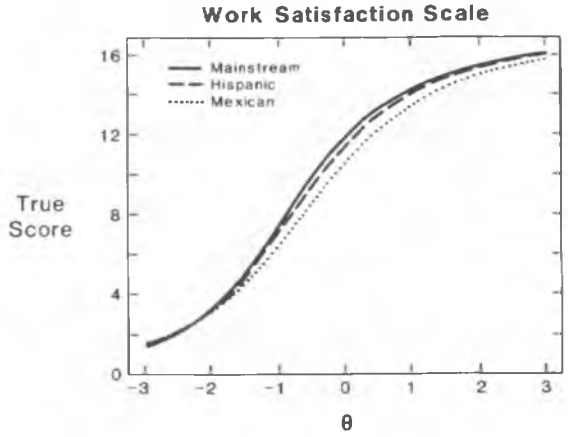

Figure 3. Test characteristic curves for the Work Satisfaction scale: U.S. mainstream, U.S. Hispanic, and Mexican samples.

REFERENCES

Birnbaum A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord & M.R. Novick, *Statistical theories of mental test scores* (pp. 395-479). Reading, MA: Addison-Wesley.

Bock, R.D., & Lieberman, M. (1970). Fitting a response model for *n* dichotomously scored items. *Psychometrika, 35*, 179–197.

Campbell, J.P. (1970). The way it should be done. *Contemporary Psychology, 15*, 540–542.

Drasgow, F. (1982). Biased test items and differential validity. *Psychological Bulletin, 92*, 526-531.

Drasgow, F. (1984). Scrutinizing psychological tests: Measurement equivalence and equivalent relations with external variables are the central issues. *Psychological Bulletin, 95*, 134–135.

Drasgow, F. (1987). A study of the measurement bias of two standardized psychological tests. *Journal of Applied Psychology, 72*, 19–29.

Drasgow F. (1988). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement*, accepted for publication.

Drasgow, F., & Kang, T. (1984). Statistical power of differential validity and differential prediction analyses for detecting measurement nonequivalence. *Journal of Applied Psychology, 68*, 498–508.

Holland, P.W., & Thayer, D.T. (1986). *Differential item performance and the Mantel–Haenszel procedure*. Paper presented at the American Educational Research Association Annual Meeting, San Francisco, CA.

Hulin, C.L. (1987). A psychometric theory of evaluations of item and scale translations: Fidelity across languages. *Journal of Cross-Cultural Psychology, 18*, 115–142.

Hulin, C.L., Drasgow, F., & Komocar, J. (1982). Applications of item response theory to analysis of attitude scale translations. *Journal of Applied Psychology, 67*, 818-825.

Hulin, C.L., Drasgow, F., & Parsons, C.K. (1983). *Item response theory: Application to psychological measurement*. Homewood, IL: Dow Jones-Irwin.

Lim, R. (1987). *Detection of biased items with marginal maximum likelihood and Bayesian estimates of item parameters*. Masters thesis, University of Illinois at Urbana-Champaign.

Lord, F.M. (1952). A theory of test scores. *Psychometric Monograph*, No. 7.

Lord, F.M. (1968). An analysis of the Verbal Scholastic Aptitude Test using Birnbaum's three-parameter logistic model. *Educational and Psychological Measurement, 28*, 989–1020.

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* (pp. 19-29). Amsterdam: Swets and Zeitlinger.

Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.

McLaughlin, M.E., & Drasgow, F. (1987). Lord's chi-square test of item bias with estimated and with known person parameters. *Applied Psychological Measurement, 11*, 161–173.

Mislevy, R.J., & Bock, R.D. (1985). Implementation of the EM algorithm in the estimation of item parameters: The BILOG computer program. In D.J. Weiss (Ed.), *Proceedings of the 1982 Item Response Theory/Computerized Adaptive Testing Conference* (pp. 189-202). Minneapolis: University of Minnesota, Department of Psychology, Computerized Adaptive Testing Laboratory.

Robinson, J.P., Athanasiou, R., & Head, K.B. (1968). *Measures of occupational characteristics*. Ann Arbor, MI: ISR.

Segall, D. (1983). *Assessment and comparison of techniques for transforming parameters to a common metric in item response theory*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Smith, P.C., Kendall, L., & Hulin, C.L. (1969). *The measurement of satisfaction in work and retirement*. Chicago: Rand McNally.

Stocking, M.L., & Lord, F.M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*, 201-210.

Thissen, D., Steinberg, L., & Gerrard, M. (1986). Beyond group-mean differences: The concept of item bias. *Psychological Bulletin, 99*, 118–128.

Vroom, V.H. (1964). *Work and motivation*. New York: Wiley.

Wood, R.L., Wingersky, M.S., & Lord, F.M. (1976). *LOGIST: A computer program for estimating examinee ability and item characteristic curve parameters* (Report No. RM 76-6). Princeton, NJ: Educational Testing Service.

Received October 20, 1986
Revision received July 24, 1987

AUTHORS

FRITZ DRASGOW. *Institutional affiliation:* University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A. *Title:* Associate Professor. *Degree:* Ph.D., 1978, University of Illinois. *Specializations:* Applied psychometrics; industrial-organizational psychology; measurement of culturally diverse groups. *Mailing address:* Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, U.S.A.

CHARLES HULIN. *Institutional affiliation:* University of Illinois at Urbana-Champaign, Champaign, Illinois, U.S.A. *Title:* Professor of Psychology. *Degree:* Ph.D., 1963, Cornell University, U.S.A. *Specializations:* Industrial-organizational psychology; measurement. *Mailing address:* Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, U.S.A.

RESUMEN EXTENSO

Medición Transcultural

FRITZ DRASGOW

y

CHARLES L. HULIN

En este artículo se describe un método para determinar si se pueden realizar comparaciones que tengan sentido entre las puntuaciones de una escala de actitudes de un grupo cultural que habla un idioma y de otro grupo cultural que habla un idioma diferente. Hay dos aspectos importantes. Primero, una escala debe presentar *equivalencia en la medición* para poder comparar las puntuaciones. Existe equivalencia en la medición si los individuos con igual posición en el rasgo medido, pero provenientes de culturas diferentes, tienen las misma puntuación observada según anticipada. Si una escala permite la medición equivalente, entonces el investigador puede estudiar si el instrumento tiene relaciones equivalentes con variables externas a través de las culturas.

En este trabajo se presentan tres métodos para examinar la equivalencia en la medición. El primero, llamado validez diferencial, examina la igualdad de las correlaciones entre la escala y una segunda variable a través de las culturas. Se ha demonstrado que este método carece de valor por su poder limitado. El segundo enfoque es el de predicción diferencial. En éste las regresiones de una segunda variable en la escala son comparadas a través de las culturas. Se ha demostrado que la falta de equivalencia en la medición conduce a ecuaciones de regresión desiguales. Sin embargo, el análisis de predicción diferencial confunde los aspectos de equivalencia de medida y relaciones equivalentes con variables externas. Las diferencias observadas en las ecuaciones de regresión pueden ser el producto de cualquiera de estas fuentes.

El tercer tipo de análisis examina la impureza del item mediante la teoría de respuesta de item. Este *no* confunde la evaluación de la equivalencia de medición con la evaluación de relaciones equivalentes con variables externas. Más aun, éste conlleva una forma estadística poderosa para identificar la falta de equivalencia en la medición.

El enfoque de la teoría de respuesta de item es ilustrado mediante un análisis de la traducción al español de una escala de actitud hacia el trabajo, el *Job Descriptive Index* (JDI; Smith, Kendall, & Hulin, 1969). Tres grupos de personas, todos empleados de la misma organización internacional de mercadeo al detal, contestaron el instrumento (JDI) ya fuera en su versión original en inglés o la traducción al español. Las tres muestras consistieron de 169 hispano-estadounidenses empleados en la ciudad de Nueva York y en Miami, 486 norteamericanos (presuntamente monolingües) y 346 mexicanos, en su mayoría monolingües, empleados en Ciudad México.

En general, hubo relativamente pocas diferencias (11 de 68 items) al comparar los norteamericanos y los hispano-estadounidenses. Al comparar los norteamericanos y los mexicanos se encontró 21 diferencias significativas y entre los hispano-estadounidenses y mexicanos hubo 9 diferencias significativas. Estos resultados difieren de un

estudio previo (Hulin, Drasgow, & Komocar, 1982) en el cual se encontraron solamente tres diferencias significativas usando la misma traducción del instrumento con un grupo de bilingües. Por lo tanto, las evaluaciones de traducciones que toman como fundamento las respuestas de personas bilingües son problemáticas. Finalmente, los patrones de diferencias significativas entre los items ofrecen evidencia sobre diversas fuentes (lingüísticas versus culturales) para la falta de equivalencia en la medición.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp. 25-40

The Effects of Generative Activity, Orienting Instructions, and Precision of Elaborations on Memorability and Metacognitions

LUCY DURGADEEN

and

FRANCIS J. DI VESTA

Pennsylvania State University
U.S.A.

This paper describes a study of the elaborative activities of successful and less successful students. The findings of the study showed that successful students generated significantly more precise elaborations than less successful students. Moreover, the precise elaborations facilitated recall even when participants were unaware of the effects of precision of elaboration on retention. The use of instructions orienting students to elaborate base sentences significantly facilitated comprehension and recall. The success of these orientations indicated that some students with a production deficiency had the required knowledge and skills but failed to employ strategies spontaneously.

Elaborations facilitate learning and enhance memory (Anderson & Reder, 1979; Bransford, Franks, Morris, & Stein, 1979; Craik & Tulving, 1979). Through elaborative activities learners make associations to target (to-be-learned) concepts and relate this new information to prior knowledge in complex ways.

Elaborations that mainly increase the complexity of a sentence will fail to enhance recall. Effective elaborations depend on such factors as availability of prior knowledge, ability to identify and form nonarbitrary relationships, and practice in making such relationships. Effective elaborative activities facilitate the formation of explanations based on structural rules, organizations, and patterns. For example, effective elabora-

Copyright © 1987 by Lucy Durgadeen and Francis J. Di Vesta. All rights reserved.

tions of the principle of conservation of energy or of Newton's laws of motion would be expected to facilitate the student's ability to explain and apply the laws or principles. Rote memorization of the laws would lead only to superficial paraphrase or restatements with no transfer value. These superficial processes are characteristic of novices when first exposed to scientific laws and principles. They tend to learn the information arbitrarily (Stein & Bransford, 1979).

The study of elaborative activities appears under two research frameworks: the levels of processing model and schema theory.

THE LEVELS OF PROCESSING MODEL AND ELABORATIONS

An underlying assumption in the levels of processing model is similar to one in perception: Stimuli are rapidly analyzed at different levels or stages. (It should be recognized that levels of processing in this article refers to degree of meaningful processing for given purposes and thereby departs from the original levels of processing model that equated semantic processing with deep processing and acoustic processing with shallow processing.) The durability of a trace and its retrievability are functions of the depth to which information processing takes place. Elaboration of information contributes to the formation of a durable trace by requiring the learner to activate essential knowledge relevant to the target information and creating relational links between the new and old information. Paris and Lindauer (1977) claimed that "the enhancement in retention that is observed in studies of the elaboration phenomenon can be attributed to the deeper level of processing" (p. 48).

Still within a levels of processing framework other researchers have shown that a change in the number and type of elaborations affects depth of processing. As Anderson and Reder (1979) indicate in the following quotation, subjects store, along with the target words, other pieces of related information:

> There is probably a purpose for the large amount of elaboration that is generated. We speculate that the rich elaboration affects memory performance. That is, we take the depth-of-processing results as an indication of a function and a consequence of the elaborative process—improved memory for material elaborated. (p. 387)

Without precise elaborations an idea to be encoded remains as a relatively isolated "node" of information and its retrieval will appear to be fragile because it has only a few routes within the cognitive network by which it can be accessed and activated. Thus, an elaborated idea is more easily assessed than one with no interrelationships. Elaborations supported by deep rules and numbers of rules aid the formation of distinct, discriminable traces and enhance performance in a retrieval environment (Jacoby & Craik, 1979).

The research on elaborative activities within a level of processing framework provides findings demonstrating that elaborations facilitate deep processing. However, clarification of the conditions under which an elaboration is effective for learning and retrieving information places the study of elaborations in a schema theory framework.

SCHEMA THEORY

The concept of a schema is frequently employed to describe the organization of knowledge. Attempts to represent the organization of knowledge range from a hierarchical and ordered arrangement (Collins & Quillian, 1969) to a more diffuse arrangement reflecting family resemblances (Wittgenstein, 1953). However, such attempts to represent knowledge reflect inadequately the complexity of knowledge that an individual possesses. The concept of a schema attempts to account for both the structure and dynamic nature of frameworks representing a person's knowledge. Consider the abstract nature of a schema in the following definitions:

> Schemata are abstract structures that represent what one holds to be generally true of the world. (Schallert, 1982, p. 20)
>
> A schema . . . is a data structure for representing the generic concepts stored in memory. (Rumelhart, 1980, p. 34)
>
> Schemata represent the generic concepts underlying objects, events, and actions. Schemata are abstract in the sense that they contain a "variable," "slot," or "place holder" for each constituent element in the knowledge structure. An important aspect of a schema is the specification of the network of relations that hold among the constituents. (Anderson, Reynolds, Schallert, & Goetz, 1977, p. 369)

A schema consists of numerous ideas connected by relational links (J.R. Anderson, 1985). Depending on how often they have been used, these links can be strong or weak. A particular concept or idea can be activated by contexts, priming, or similar events. Further, once activated the idea spreads to other ideas or concepts connected to the one activated. A study of elaborations in a schema theory framework focuses on the nature of the elaborations and the type of linkages that the elaborations allow.

The significance of precise linkages or relationships is evident in encoding textual information. For example, in order to read and comprehend a text, a person must make inferential elaborations (Thorndyke, 1976). Explicit texts require fewer inferences than implicit texts. The reader builds an implicit text by elaborating the given information in order to create a context within which comprehension is possible. Difficulties in comprehending implicit texts might be due to an inability of a learner to elaborate the information to make it meaningful. Practice in the selection and use of appropriate prior knowledge in learning new information helps in the formation of

meaningful relationships (Anderson & Reder, 1979; Franks, Vye, Auble, Mezynski, Perfetto, Bransford, Stein, & Littlefield, 1982; Stein & Bransford, 1979). Precise elaborations function to make what are initially arbitrary relations more understandable. They provide the links between the base source domain (prior knowledge) and the target domain (new information) that reduce the arbitrariness of the to-be-learned information, as illustrated in subsequent paragraphs.

The importance of precise elaborations is also evident at the sentential level. Consider the basic sentence:

The tall man bought the crackers.

Subjects received a set of such sentences, each dealing with a different type of man (Stein, Littlefield, Bransford, & Persampieri, 1984). Ability to recall in a cued recall test the adjective describing the man in each sentence depended on the nature of the elaboration of the base sentence. A precise elaboration of the basic sentence, "The tall man bought the crackers *which were on the top shelf*," facilitated recall. An imprecise elaboration, "The tall man bought the crackers *which were on sale*," failed to help subjects discriminate between types of men on a later recall task (e.g., "Which man bought the crackers?").

The precise elaboration exemplified above represents a nonarbitrary relationship between the subject "tall man" and "top shelf." Tall people logically can reach top shelves easier than short people. The imprecise elaboration represents an arbitrary relationship since the target object, that is, the characteristics of the person who might have bought crackers on sale, is unconstrained. Thus, the relationship provides no basis for a meaningful relation or explanation.

Studies of a student's ability to generate and use precise relationships show that successful students more than less successful students are aware of the powerful effect of precision of elaboration on learning and retrieving information. Stein, Bransford, and Owings (1982) found that successful students when given base sentences, generated significantly more precise elaborations than less successful students. The importance of elaborations in cognitive processes was demonstrated when less successful students who received training in elaborative activities improved in both acquisition and unintentional memory tasks. The students learned to select and activate prior knowledge appropriate for the target information. They also learned how to generate precise elaborations to link new and old information. And they evaluated the effect of making precise elaborations on their own learning and memory performance.

The formation of precise elaborations is possible through several learning activities including paraphrasing, summarizing, creating analogies, generative note taking, and question answering. Imaginal narrative elaboration (Bower, 1970; Bower & Clark, 1969; Paivio, 1971; Rohwer, 1966) also facilitates learning. However, the effectiveness of an elaborative strategy ultimately depends on actively engaging the student in forming the elaborations as a generative process (Wittrock, 1974).

The aforegoing rationale guided the design of one phase of the present study which examined the elaborative activities of successful and less successful students (Durgadeen, 1986). In another phase, the techniques employed in other studies (e.g., Di

Vesta & Finke, 1985; Stein, Littlefield, Bransford, & Persampieri, 1984) were modified to provide applications to simple concepts in general science. In a third phase, we examined metacognitions accompanying elaboration activities. Our major concern, then, was to determine the potential application of elaboration as a teaching/learning strategy in classroom settings.

METHOD

Participants

The participants were 88 second-year students in a secondary school in San Fernando, Trinidad: 47 of whom were boys and 41 were girls. They were randomly assigned to three groups for the three phases of the study. The age range of the students selected for the study was 12.5 to 14.1 years, with a mean chronological age of 13.2 years.

The classification of the students as successful and less successful was made by ranking all of the students in the second year on the basis of scores on teacher-made science tests. The test consisted of 50 multiple-choice items and 5 essay questions (10 points each). Both types of question tested mainly recall of information. The top 44 students were the successful (knowledgeable) group. Their scores ranged from 86 to 69 ($M = 73.00$, $SD = 6.05$). The bottom 44 were the less successful (less knowledgeable) group. Their scores ranges from 59 to 30 ($M = 49.23$, $SD = 7.94$). Classroom teachers examined the groupings and used the students' academic records to verify the classifications. No exceptions to the original classifications of these additional criteria are required.

Materials

The most important variable for the present study consisted of three sentence types: base sentences, precisely elaborated base sentences, and imprecisely elaborated base sentences. A list of 15 sentences made up the acquisition task. The statements in each set were equally divided into base ($n = 5$), precisely elaborated ($n = 5$), and imprecisely elaborated ($n = 5$) statements. They were presented in random order.

The base sentences contained two propositions with water as the subject. All the base sentences had the same frame: The _____ water was in a _____. For example:

The impure water was in a vial.

The soapy water was in a basin.

The imprecise elaborations contained the base sentence and a sentence ending that formed an arbitrary relationship with the target concept — water. In other words, it did

not provide a meaningful link with the target concept that would reduce confusion among the types of water. For example:

The impure water was in a vial that was kept in the store room.

The soapy water was in a basin that was made of plastic.

The precise elaborations contained the base sentence and a sentence ending that had a nonarbitrary relationship with the rest of the sentence. They were specific, in an explanatory sense, to the extent that they explained why a certain kind of water was in a given container. For example:

The impure water was in a vial *for testing in the lab.*

The soapy water was in a basin *for washing dishes.*

Recall Test

The recall task tested recall of the adjective, as the target word, in each sentence. There were 15 questions corresponding to the sentences in the acquisition set. All questions were based on the frame: Which water was in a(n) _____? For example: Which water was in a vial?

Which water was in a basin?

In addition to answering the questions, the participants rated their certainty about the correctness of their answers on a scale of 1 (very certain answer is correct) to 5 (very uncertain answer is correct).

Procedure

The overall study consisted of three phases (or three separate, but related, studies).

Phase 1. In Phase 1, 20 students—10 successful and 10 less successful—were tested individually in the school library to determine their ability to generate meaningful elaborations and thereby examine the effect of such generation on recall.

The materials used were 8 base sentences. The students were told that the purpose of the study was to find out why some sentences were easier to comprehend and remember than others. They were then given the eight base sentences on a typewritten sheet and were asked to construct, for each sentence, a sentence ending that would help them to remember the sentence to be recalled later. Two minutes of interpolated activity followed the elaboration generation exercise to reduce effects of retention in short-term memory. During this time the students discussed the second-year program at their school. The cued recall test was then administered. Each participant was allowed 30 seconds to write a response to each question. The elaborations generated were independently classified as precise or imprecise by three judges who were trained to make these decisions. There was complete agreement among the judges since precise elaborations were easily identified.

Phase 2. The participants in Phase 2, different from those in Phase 1 but from the same pool, were 48 students: 24 successful and 24 less successful. They, too, were tested individually. Half of both groups received relevant orienting instructions, the other half irrelevant orienting instructions (defined in the next paragraphs).

Sentence sets. These variables were crossed with a third between-subjects variable, sentence-set (hereafter, referred to as A, B, or C). Each sentence-set contained the same base sentences that might have appeared as one sentence-type in one set, but as another sentence-type in another set. Thus, in one set the base sentence might be presented, in the other set the base sentence might be imprecisely elaborated, and in the third set it might have been precisely elaborated. Items for a given set were randomly selected from a pool of items with the restriction that no base item would appear twice in a sentence set. All students were told that the exercise was an attempt to find out why some sentences were more difficult to understand and remember than others. They were told to listen to sentences that the researcher would read only once. After each sentence they were to use a comprehensibility rating scale to rate how easy or difficult the sentence was to comprehend. They were also told that a recall test would follow the comprehensibility rating exercise.

Orienting instructions. Those students who were administered *relevant* orienting instructions were given examples of base, precisely elaborated, and imprecisely elaborated sentences. They were asked to attend to the precisely elaborated sentence and to the fact that it provided a reason for putting a certain type of water in a given container. Those students who received *irrelevant* orienting instructions were told to attend to the sentence and to the fact that some were longer than other sentences in describing the event.

After each sentence was read, the students used a rating scale from 1 (completely easy to understand) to 10 (completely difficult to understand) for rating the comprehensibility of the sentence. Two minutes of interpolated activity followed, after which the cued recall test was administered.

Phase 3. In Phase 3, 20 different students from the original pool of 88 students — 10 successful and 10 less successful — were administered a rating task to examine their metacognitions about memorability of different sentential forms. The students were tested in groups of five. Each student received a difficulty rating sheet that contained five sets of sentences. Each set contained a base sentence, and a precise and an imprecise elaboration of the sentence.

Example: The short man used the broom. (base)
The short man used the boom to *sweep the porch*. (imprecise)
The short man used the broom *to operate the light switch*. (precise)

The students were asked to use a scale from 1 (very easy) to 5 (very difficult) to rate the sentences as easy or difficult to remember. They also gave a reason for their choice of easiest and most difficult sentence in each set of sentences.

RESULTS

Phase 1

The number of precise elaborations generated by successful and unsuccessful groups was compared via Student's t test. The results of this analysis showed that the difference in the number of precise elaborations generated by the two groups was significantly different, $t(18) = 2.67, p < .05$. This result indicated that the number of precise elaborations generated by successful students ($M = 5.00, SD = 4.00$) was almost twice that generated by the less successful students ($M = 2.8, SD = 2.4$). Further, the correlation between precision of elaboration and recall scores was .69 indicating that the more precise elaborations favorably influenced memorability.

Phase 2

The data for Phase 2 were analyzed by a 2 x 2 x 3 x 3 mixed analysis of variance. The between-subjects factors were student level (successful and less successful students), type of orienting instruction (relevant and irrelevant), and sentence-set (lists A, B, and C). (The reader will recall that sentence-set refers to three different lists of sentences selected at random from a pool of sentences. Within each set were represented the same number of base, precisely elaborated, and imprecisely elaborated sentences.) The within-subjects variable was sentence type (base sentence, precisely elaborated sentences, and imprecisely elaborated sentences).

The results of the analysis are summarized in Table 1. The analysis yielded significant main effects for type of orienting instruction $F(1, 36) = 4.41, p < .05$ and for sentence type $F(2, 72) = 20.27, p < .01$. Since sentence-set could be treated as a random variable (i.e., other sentences might have been used in this study or might be used in a replication of the study) a random effects model is appropriate for the analysis of the sentence-set effects. This analysis yielded essentially the same results as the analysis via the fixed effects model, with the exception that the effect of orienting instructions yielded $F = 7.33, p < .01$. (This effect was calculated by dividing the mean square for orienting instructions by the mean square for the interaction of orienting instructions and sentence set.) Other relevant analyses involved in the random effects model yielded F ratios < 1.00.

The average recall score for participants administered relevant orienting instructions ($M = 2.06$) was significantly greater than that for students administered irrelevant orienting instructions ($M = 1.53$). The differences among the means were tested via Tukey's Wholly Significant Difference (WSD) statistic ($df = 2, 36, p < .05$), yielding a critical value of 0.51. The mean recall scores of sentences precisely elaborated ($M = 2.52$) were found in these comparisons to be significantly ($p < .05$) greater than recall

Table 1

Summary of Analysis of Variance of Recall Scores

Source	*df*	*MS*	*F*
Between subjects			
Student level (S)	1	0.11	0.05
Orienting instructions (I)	1	10.03	4.41*
S x I	1	1.00	0.44
Sentence set (G)	2	2.25	0.99
S x G	2	0.86	0.38
I x G	2	1.36	0.60
S x I x G	2	2.25	0.99
Error	36	2.27	
Within subjects			
Sentence type (T)	2	21.02	20.27*
S x T	2	0.84	0.81
I x T	2	0.49	0.05
S x I x T	2	1.40	1.35
G x T	4	4.18	4.03*
S x G x T	4	0.43	0.42
I x G x T	4	1.10	1.06
S x I x G x T	4	0.30	0.29
Error	72	1.04	

*$p < .05$

scores based on the base sentences ($M = 1.62$) and imprecisely elaborated ($M = 1.23$) sentences. The effect due to sentence-set (i.e., A, B, or C) was primarily the result of lower recall scores (overall $M = 1.40$) for sentence set C, which apparently contained at least one ambiguous item in the precisely elaborated statement, "The dam water was in a pan for the animals to drink."

Analysis of the comprehensibility ratings (see Table 2) yielded a significant main effect for sentence-type, $F(2, 292) = 5.51$, $p < .05$. Thus learners rated precisely elaborated ($M = 1.63$) sentences as significantly easier to comprehend than the imprecisely elaborated ($M = 2.42$) sentences. Although the base sentences were rated ($M = 1.89$) slightly more difficult to understand than precisely elaborated sentences, the difference was not significant ($p > .05$).

Table 2

Summary of Analysis of Variance of Comprehensibility Ratings

Source	*df*	*MS*	*F*
Between subjects			
Student level (A)	1	8.56	1.60
Error	46	1.42	
Within subjects			
Sentence type (J)	2	7.81	5.51*
A x J	2	0.40	0.28
Error	92	1.42	

*$p < .05$

Phase 3

The students' ratings of sentences as easy or difficult to remember were analyzed by a 2 x 3 mixed analysis of variance. The between-subjects factor was student level (successful and less successful students). The within-subjects factor was sentence type (base, precisely elaborated, and imprecisely elaborated). The results of the analysis are summarized in Table 3. As can be seen in that table, only the effect for sentence type was significant, $F(2, 36) = 18.16$, $p < .001$. Students rated base sentences ($M = 1.25$, $SD = 0.5$) as significantly ($p < .05$) easier to remember than imprecise elaborations ($M = 2.05$, $SD = 0.60$) and precise elaborations ($M = 2.51$, $SD = 0.84$). The data indicate

Table 3

Summary of Analysis Variance of Difficulty Ratings of Sentences by Successful and Unsuccessful Science Students

Source	*df*	*MS*	*F*
Between subjects			
Student level (S)	1	0.82	1.82
Error	18	0.44	
Within subjects			
Sentence type (T)	2	8.13	18.16*
S x T	2	0.12	0.26
Error	36	0.45	

*$p < .001$

that memorability is improved by imposing precise elaborations on text statements. Nevertheless, the students do not possess the metacognitions about the value of precisely elaborated statements for retrieval.

DISCUSSION

The study was different from those conducted earlier (Bransford, Franks, Morris, & Stein. 1979; Bransford, J.D., Stein, N.J., Vye, N.J., Franks, J.J., Auble, P.M., Mezynski, K.J., & Perfetto, G.A., 1982; Stein et al., 1982) in two ways: First the study was conducted with secondary school students in a Caribbean culture (Trinidad), a culture quite different from the learning environment of students in all of the earlier studies. Secondly, the materials were based on facts that are typically part of knowledge gained from a subject matter domain taught at the lower secondary grades. The orientation in the study was that if information is to be understood it must be linked either in the text or by the learner to information already in the learner's cognitive structure. Thus, the implication is that basic information must be explained; mere presentation of facts is not an optimal learning/teaching procedure.

The use of sentences related to simple facts in science was a departure from arbitrary base sentences used in earlier studies, most of which were similar to those employed by Stein and Bransford (1979). Accordingly, the present investigation implies the potential of further extending the methods of the present study to text-like *passages*, the sentences of which can be elaborated in ways to provide differential encodings.

The results obtained in Phase 1 showed that when given base sentences successful students do autonomously generate significantly more precise elaborations than do less successful students. Further, the independent generation of precise elaborations facilitated their recall.

Since the elaboration of sentences appears to have the potential of being an easily taught strategy, these results imply that teaching such strategies would significantly improve understanding. Students in Phase 2 who received relevant orienting instructions (i.e., to attend to the relationships in precisely elaborated sentences) presumably processed the information more deeply and, as a consequence, recalled significantly more items than those students who were administered irrelevant orienting instructions (i.e., to attend to differences in length of sentences). Relevant instructions can be regarded as one means of making students aware of the importance of elaborative activities in learning. It also encourages "deep" processing. The finding that learning was enhanced by this training, limited though it was, has implications for incorporating simple orienting devices for learning-how-to-learn strategies into regular school teaching activities.

The recall scores for precisely elaborated sentences in both Phases 1 and 2 were significantly greater than those from imprecise elaborations and base sentences (short sentences) despite the fact that students considered base sentences as more comprehensible than either precisely or imprecisely elaborated sentences. These findings clearly

imply the robustness of the effect. They also imply that although students may be unaware of factors (such as precise elaborations) that affect ease of learning, they can and do benefit from precise elaborations build into the learning tasks. This finding has implications for authors of texts and curriculum materials. Properly constructed explanations (elaborations) should help students bridge old and new information when the students lack the metacognitive skills (that is, when there is a mediational or production deficiency, or a production inefficiency—see Flavell, 1970) to generate appropriate elaborations.

Finally, a note should be made regarding the three sentence sets. Although each set has been pretested and found reliable with U.S. students, there was not a direct correspondence when used in Trinidad. Items (in Set C) that were precisely elaborated for U.S. students were ambiguous for Trinidad students. Further studies are needed to determine the loci of such differences.

In summary, the results of the study clearly demonstrated that a characteristic of successful (knowledgeable) students in a content area is their ability to generate significantly more precise elaborations than less successful (knowledgeable) students. Such spontaneous generation of precise elaborations is characteristic of knowledgeable persons such as experts. Further, both successful and unsuccessful students benefited from training even when it was limited to brief orienting instructions. Students also benefited from precise elaborations when retrieving information even though they were unaware of the influence of elaboration on memorability.

An important concern in this study, as in its earlier counterparts, is that the conditions under which the lists were constructed and administered make learning and recall of base sentences especially subject to massive interference (all sentences dealt with water). Yet by use of precise elaborations the effects of interference were effectively minimized, as shown in the recall scores.

The findings in this study demonstrate the robust effect of elaborating as a learning strategy, not as a matter of simply extending the information in arbitrary ways but as a means of explaining information so that its structure is made logical to the learner. The findings of the study hold promise for simple techniques by which to improve student performance through teaching students how to select and use learning strategies to bridge old and new information in precise ways, a procedure that requires little additional expenditure of teaching time and is easily incorporated into typical curricular contexts.

REFERENCES

Anderson, J.R. (1985). *Cognitive psychology and its implications*. San Francisco: W.H. Freeman.

Anderson, J.R., & Reder, L.M. (1979). An elaborative processing explanation of depth of processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 385–403). Hillsdale, NJ: Erlbaum.

Anderson, R.C., Reynolds, R.E., Schallert, D.L., & Goetz, E.T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal, 14*, 367–381.

Bower, G.H. (1970). Analysis of a mnemonic device. *American Scientist, 58*, 496–510.

Bower, G.H., & Clark, M.C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science, 14*, 141–182.

Bransford, J.D., Franks, J.J., Morris, C.D., & Stein, B.S. (1979). Some general constraints on learning and memory research. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 331–354). Hillsdale, NJ: Erlbaum.

Bransford, J.D., Stein, N.J., Vye, N.J., Franks, J.J., Auble, P.M., Mezynski, K.J., & Perfetto, G.A. (1982). Differences in approaches to learning: An overview. *Journal of Experimental Psychology, 111*, 390–398.

Collins, A.M., & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 240–247.

Craik, F.I.M., & Tulving, E. (1979). Depth of processing and retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268–294.

Di Vesta, F.J., & Finke, F.M. (1985). Metacognition, elaboration, and knowledge acquisition: Implications for instructional design. *Educational Communications and Technology Journal, 33*, 285–293.

Durgadeen, L. (1986). *Elaborations: Differences in the metacognitive behavior of successful and less successful students*. Unpublished doctoral dissertation, Pennsylvania State University, University Park.

Flavell, J.H. (1970). Developmental studies of mediated memory. In H.W. Reese & L.P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 5, pp. 182–211). New York: Academic.

Franks, J.J., Vye, N.J., Auble, P.M., Mezynski, K.J., Perfetto, G.A., Bransford, J.D., Stein, B.S., & Littlefield, J. (1982). Learning from explicit versus implicit texts. *Journal of Experimental Psychology: General, 111*, 414–422.

Jacoby, L.L. & Craik, F.I.M. (1979). Theoretical and empirical developments. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 1–21). Hillsdale, NJ: Erlbaum.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Paris, S.G. & Lindauer, B.K. (1977). Constructive aspects of children's comprehension and memory. In R.K. Kail, Jr. & J.W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 35–60). Hillsdale, NJ: Erlbaum.

Rohwer, W.D., Jr. (1966). Constraints, syntax and meaning in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior, 5*, 541–547.

Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Erlbaum.

Schallert, D.L. (1982). The significance of knowledge: A synthesis of research related to schema theory. In W. Otto & S. White (Eds.), *Reading expository material* (pp. 13–48). New York: Academic.

Stein, B.S., & Bransford, J.D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning and Verbal Behavior, 18*, 769–777.

Stein, B.S., Bransford, J.D., & Owings, R.A. (1982). Differences in the precision of generated elaborations. *Journal of Experimental Psychology: General, 111*(4), 406–413.

Stein, B.S., Littlefield, J., Bransford, J.D., & Persampieri, M. (1984). Elaboration and knowledge acquisition. *Memory and Cognition, 12*, 522–529.

Thorndyke, P.W. (1976). The role of inferences in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 15*, 437–466.

Wittgenstein, L. (1953). *Philosophical investigations*. (G.E. Anscombe, Trans.). New York: Macmillan.

Wittrock, M.C. (1974). Learning as a generative process. *Educational Psychologist, 11*, 87–95.

Received November 17, 1986
Revision received July 29, 1987

AUTHORS

LUCY DURGADEEN. *Institutional affiliation:* Ministry of Education, Port-of-Spain, Trinidad, West Indies. *Title:* Science Curriculum Officer. *Degree:* Ph.D., 1985, Pennsylvania State University, U.S.A. *Specialization:* Curriculum and instruction. *Mailing address:* 305 Soogrim Street, Gulf View, La Romain, Trinidad, West Indies.

FRANCIS J. DI VESTA. *Institutional affiliation:* Pennsylvania State University, U.S.A. *Title:* Professor of Education and Psychology. *Degree:* Ph.D., 1949, Cornell University, U.S.A. *Specialization:* Learning, memory, and cognition. *Mailing address:* 314 Cedar Building, Division of Educational and Counseling Psychology, Pennsylvania State University, University Park, PA 16802, U.S.A.

RESUMEN EXTENSO

Efectos de la Actividad Generadora, las Instrucciones Orientadoras y la Precisión de las Elaboraciones en la Memoria y la Metacognición

LUCY DURGADEEN

y

FRANCIS J. DI VESTA

Se investigaron los efectos de las elaboraciones en la capacidad para memorizar información sencilla relacionada con la ciencia. Los participantes fueron niños de Trinidad y Tobago en el segundo año de escuela secundaria. Se asumía que la elaboración precisa llevaba a un procesamiento más profundo de la idea objetivo que una frase sin elaborar o una frase que se extendiera meramente hacia una expresión más compleja. Las elaboraciones precisas representan relaciones no arbitrarias en la información que debe ser aprendida proveyendo de ese modo claridad a la información de interés. Como consecuencia, se asume que las elaboraciones precisas facilitan el aprendizaje y la memoria.

Las formas de elaboración pudieran ser ilustradas como sigue: (a) Frases base sin elaborar que manifiestan solamente relaciones arbitrarias, ej., *El agua de desecho estaba en una cañería.* (b) Elaboraciones imprecisas que extienden las frases de modo que no se relacionan con el objetivo de la información (la clase de agua), ej., *El agua de desecho estaba en una cañería que tenía tres secciones.* (c) Elaboraciones precisas que relacionan la información de interés con una actividad relevante, ej., *El agua de desecho estaba en una cañería que conducía a la planta de filtro.*

Tres estudios se llevaron a cabo. En el primero, examinamos hasta qué punto estudiantes con éxito en ciencias ($n = 10$) y estudiantes sin éxito en ciencias ($n = 10$) generaban elaboraciones precisas de las frases base. En el segundo estudio ($n = 48$) los participantes fueron asignados por igual a los grupos de tratamiento compuestos por tres variables entre sujetos, cruzadas ortogonalmente: éxito de los estudiantes en ciencias (24 con alta puntuación y 24 con baja puntuación en un examen de ciencias), instrucciones (relevantes o irrelevantes) sobre el uso de elaboraciones y grupos de frases (tres listas de frases distintas). La variable intra sujetos fue la clase de elaboración (base, precisa e imprecisa). Igual número ($n = 5$) de cada uno de los tres tipos de frases fue incorporado en cada lista de 15 frases sobre las distintas clases de agua. Dependiendo del tratamiento al que fueron asignados los participantes, se les leyó una de las listas. La variable dependiente fue el número de frases base, imprecisamente elaboradas o precisamente elaboradas que fueron recordadas. En el tercer estudio, las metacogniciones de los estudiantes con éxito ($n = 10$) y sin éxito ($n = 10$) fueron obtenidas con las evaluaciones en la facilidad para aprender las tres clases de frases elaboradas.

En el primer estudio, los estudiantes con éxito autogeneraron significativamente más elaboraciones precisas que los estudiantes con menos éxito. La precisión de elaboración se correlacionó con el recuerdo. En el segundo estudio, el efecto de instrucciones relevantes fue robusto, indicando que los participantes pueden usar con mayor eficacia el conocimiento necesario y la habilidad con respecto a la elaboración, cuando son instruídos para ello. En el tercer estudio, la frase base (la frase más corta) fue evaluada como la más fácil de las tres variaciones para recordar. Así, las metacogniciones de los estudiantes no correspondieron a su actual ejecución ya que el número medio de frases precisamente elaboradas que se recordaron fue en realidad siempre significativamente mayor que el número medio de frases imprecisamente elaboradas.

El presente estudio apoya previos hallazgos con respecto al poder de la elaboración precisa en el aprendizaje y la memoria, incluso con material donde existe la oportunidad de masiva interferencia. Dado que los datos fueron obtenidos en una cultura diferente y estuvieron basados en materiales distintos a los usados en previos estudios, parece sugerirse que existe una considerable promesa en extender la presente perspectiva de elaboración hacia objetivos con materias más complejas y hacia estudios evolutivos para determinar las etapas en la habilidad para utilizar la elaboración como una actividad espontánea.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp. 41-55

Efectos a Largo Plazo (30 Meses) en la Modificación de la Impulsividad: Su Relevancia Para la Comprensión de los Estilos.

PEDRO SOLÍS-CÁMARA R.

y

PEDRO SOLÍS-CÁMARA V.

Unidad de Investigación Biomédica de Occidente,

Instituto Mexicano del Seguro Social
MÉXICO

El propósito de este estudio retrospectivo con seguimiento fue conocer si la reflexión (latencias largas y pocos errores en la prueba Matching Familiar Figures, MFF) mostrada por niños impulsivos (n = 13) entrenados para modificar su impulsividad (Solís-Cámara, 1985) persistía a largo plazo (30 meses). Un segundo propósito fue conocer si un grupo de niños reflexivos (n = 10) y uno de impulsivos (con latencias cortas y muchos errores; n = 7) no entrenados 30 meses antes, mostraban la tendencia a hacerse rápidos-exactos o reflexivos, respectivamente. Los análisis de latencias y errores en la prueba MFF por grupos no estuvieron de acuerdo a lo esperado según la literatura de reflexión–impulsividad (R–I). En forma retrospectiva se aplicó un modelo probabilístico que establece que 12 o más errores en la prueba MFF indican que la tarea fue contestada al azar (Solís-Cámara & Solís-Cámara, 1987). Los resultados indicaron que cuando los sujetos son clasificados por estilos de R–I y los sujetos que responden al azar a la prueba MFF no son incluídos, la mayoría de los niños conservaron su tiempo lento o rápido en un plazo de 30 meses, hayan o no sido tratados para convertirse a reflexivos. De acuerdo a este modelo se discute la importancia de distinguir entre capacidad y preferencia para la mejor comprensión de los estilos de R–I.

Copyright © 1987 by Pedro Solís-Cámara R. and Pedro Solís-Cámara V. All rights reserved.

En el estudio de las diferencias individuales una de las dimensiones cognoscitivas que ha sido de interés constante de parte de los investigadores ha sido la de reflexión–impulsividad. Kagan y colaboradores (Kagan, Rossman, Day, Albert & Phillips, 1964) definieron la reflexión (R) como la tendencia de unos niños a evaluar la validez de varias alternativas para elegir una solución en problemas con un número fijo de alternativas; así también, definieron la impulsividad (I) como la tendencia de otros niños a seleccionar una solución sin detenerse a evaluar varias alternativas posibles, en pruebas con gran incertidumbre en cuanto a la respuesta correcta.

La dimensión de R–I es conocida también como tiempo conceptual, ya que las respuestas del individuo representan el tiempo de decisión para evaluar o no cada respuesta alternativa y seleccionar una de ellas. Sin embargo, la operacionalización del tiempo conceptual fue realizada con una tarea de igualamiento de figuras conocidas (Matching Familiar Figures; MFF) de la cual se obtienen dos índices conocidos como latencia de la respuesta y total de errores (Kagan & Messer, 1975). El haber agregado a las latencias los errores para definir los estilos reflexivo (largas latencias con pocos errores) e impulsivo (latencias breves y muchos errores) parece ser lo responsable de los resultados contradictorios en varios estudios y de las conceptualizaciones equívocas acerca de la naturaleza de los estilos de R–I (ver Block, Block & Harrington, 1974; Kagan & Messer, 1975).

La evidencia existente, de que los estilos de R–I se relacionan con el desempeño de escolares en muy diversas tareas educativas y experimentales, así como con otros estilos cognoscitivos (ver revisiones de Ancilloti, 1985; Kogan, 1976; Messer, 1976), le ha dado popularidad a los estilos de R–I, pero ha enmascarado la validez del constructo de tiempo conceptual (Kagan et al., 1964), ya que han sido los errores en la tarea MFF los que preponderantemente se han relacionado con los errores de otras tareas. Esto ha llevado a algunos investigadores (Block, Block & Harrington, 1975; Gjerde, Block & Block, 1985) a considerar que al pretender clasificar los estilos de R–I, lo único que se está haciendo es medir la capacidad (o competencia) de los sujetos en la solución de tareas.

En nuestro propio trabajo, hemos encontrado relación entre los estilos de R–I y el desempeño en una tarea de discriminación visual sucesiva de números (Solís-Cámara, Troyo & Solís-Cámara, 1983) y hemos mostrado evidencias de la contaminación del tiempo conceptual (latencias de las respuestas) con respuestas motrices (Solís-Cámara & Solís-Cámara, 1986a) las cuales pueden indicar que las latencias de las respuestas en la tarea MFF no reflejan en algunos niños el tiempo conceptual, sino la ausencia de ciertas habilidades fundamentales. Considerando estos razonamientos y observando que la modificación del estilo impulsivo al reflexivo se ha logrado con procedimientos conceptualmente muy diferentes, Solís-Cámara mostró en un estudio reciente (1985) que el entrenamiento en un constructo fundamental como el de *atención* podía ser tan efectivo como el entrenamiento en estrategias cognoscitivas (o sea, autoinstrucciones con modelamiento en reflexión; Meichembaum & Goodman, 1971) para la modificación del estilo impulsivo. En ese estudio se mostró cómo los niños tratados por su impulsividad modificaban en forma significativa sus latencias de las respuestas y sus errores

igualando su desempeño en la tarea MFF al de un grupo de niños reflexivos que servía de control. Sin embargo, como lo escribió Solís-Cámara, la evidencia de que el entrenamiento en la habilidad para atender una tarea de discriminación visual sucesiva (de los números 1, 2 y 3) haya modificado la impulsividad presenta nuevamente interrogantes acerca de la existencia del tiempo conceptual como constructo fundamental.

Por otra parte, el aspecto evolutivo de la impulsividad hacia la reflexión puede estar relacionado con el problema de si la dimensión de R–I es reflejo de diferencias individuales (ya sea, estilos) o representa la capacidad del niño para resolver ciertas tareas. Los pocos estudios (ver Messer, 1976; Salkind, Kojima & Zelniker, 1978) sobre este tema indican que del nivel preescolar al escolar (aproximadamente a los 10 años de edad), se observa un cambio hacia la reflexión (latencias más largas con menor número de errores) pero de los 11 años en adelante las latencias decrementan y los errores incrementan ligeramente, estos últimos en forma no significativa. Estas observaciones han sido interpretadas como evidencia del cambio de un desempeño rápido-inexacto (impulsivo) a uno lento-más-exacto (reflexivo) y de ahí a uno rápido-exacto, es decir, a una mayor eficiencia (Salkind & Wright, 1977; Salkind et al., 1978). Sin embargo, todos estos estudios se infieren de comportamientos grupales y no muestran los cambios que ocurren después del paso de un tiempo prolongado en niños clasificados como impulsivos o reflexivos. Posiblemente, la ausencia de estas observaciones ha favorecido la atractiva proposición evolutiva ofrecida por esos autores en la que el niño, con la edad, se convierte de impulsivo a reflexivo y de ahí a eficaz. Sin embargo, esta proposición se ha intentado explicar de muy diversas maneras sin consolidarse, empíricamente, ninguna de ellas (Ancillotti, 1985; Messer, 1976; Salkind et al., 1978).

Los interrogantes planteados en nuestros estudios y los de otros autores (ej., Ancillotti, 1985) nos llevó a proponer un modelo probabilístico (Solís-Cámara & Solís-Cámara, 1987) que muestra que algunos de los niños designados con *estilo* impulsivo, en realidad contestan al azar (o sea, con 12 o más errores en la tarea MFF) y contaminan al resto de la muestra en la cual usualmente hay niños que no contestan al azar y que pueden estar evaluando las soluciones alternativas (Kagan, 1971) pero lo hacen rápidamente mientras que otros lo hacen lentamente.

Dado el estado actual de interrogantes con respecto a los estilos de R–I, hemos considerado que el análisis del desempeño a largo plazo (30 meses) de los niños que participaron en el estudio de modificación de la impulsividad (Solís-Cámara, 1985) podría ofrecer algunas evidencias preliminares que favorecieran la comprensión del tiempo conceptual. A pesar de que los estilos de R–I se determinan de acuerdo al procedimiento de división de las medianas, el cual ha sido justamente criticado por su imprecisión (Messer, 1976), el estudio de seguimiento del desarrollo de niños clasificados en las celdas de R–I puede ofrecer información interesante, a pesar de esta deficiencia.

Este es un estudio retrospectivo con seguimiento que tiene como objetivo conocer si dos grupos de niños clasificados con estilo impulsivo y modificados para mostrar un estilo reflexivo, con procedimientos radicalmente diferentes (habilidad para atender

vs. autoinstrucciones con modelamiento reflexivo), conservaban o modificaban nuevamente su estilo en un período relativamente largo (30 meses). Un segundo objetivo fue el de conocer si los niños clasificados como reflexivos o impulsivos y que sirvieron de grupos controles en el estudio de modificación, conservaban su estilo o si, de acuerdo a la literatura, los impulsivos tenderían a hacerse reflexivos y los reflexivos a su vez rápidos-exactos, con el paso de la edad (30 meses).

MÉTODO

Sujetos

La muestra de este estudio de seguimiento es parte de una población de 87 sujetos estudiados en un proyecto de reflexión–impulsividad llevado a cabo en la ciudad de Guadalajara (México). Las edades cronológicas de los sujetos fueron comprobadas con sus actas de nacimiento y su procedencia socioeconómica (nivel medio–bajo) fue determinada por entrevistas con la familia de cada niño (ver Solís-Cámara, 1985).

De los 38 niños clasificados como impulsivos o reflexivos en el estudio de modificación, sólo 30 fueron localizados en su escuela para este trabajo. Todos los niños incluidos ahora asistían al sexto año de escuela primaria. La remisión de sujetos se debió a que tres niños impulsivos reprobaron el quinto año escolar y cinco fueron dados de baja por razones que no se nos pudo informar. El número de sujetos de cada condición experimental original se redujo de la siguiente manera: un grupo ($n = 7$) entrenado en autoinstrucciones con modelamiento, un grupo ($n = 6$) entrenado en la habilidad para atender, un grupo control (no entrenados; $n = 7$) de sujetos impulsivos y un grupo de reflexivos control ($n = 10$), el cual fue el único que conservó su tamaño original, 30 meses antes.

De los 30 niños, 14 fueron del sexo femenino y 16 del sexo masculino. La edad cronológica media fue de 11 años 11 meses y el incremento global de edad fue de 30.5 meses (2 años y 6 meses) a partir de la primera evaluación con la prueba MFF (Solís-Cámara, 1985). El promedio de CI según Goodenough fue de 95.9 ($DE = 7.8$) y el de las calificaciones promedio fue de 8.6 ($DE = 1.05$); los datos desglosados de estas variables, para cada uno de los cuatro subgrupos considerados en el estudio, aparecen en la Tabla 1.

Procedimientos y Materiales

Tres pasantes de psicología que desconocían el propósito de este estudio y la clasificación previa de los sujetos en cuatro subgrupos, realizaron la administración en forma global de la prueba de la figura humana de Goodenough (1982) y en forma individual de la prueba MFF (Kagan & Messer, 1975); al final del ciclo escolar se recolectaron las calificaciones promedio de cada sujeto. La prueba de Goodenough y

las calificaciones escolares fueron obtenidas con el propósito de conocer si había diferencias perceptomotrices y escolares entre los niños impulsivos y los reflexivos.

RESULTADOS

Los análisis de varianza de las variables descriptivas (sexo, edad cronológica, cociente intelectual según Goodenough y la calificación promedio) no mostraron diferencias entre los grupos, excepto para el CI ($F(26) = 2.99$, $p < .05$) debido al mayor valor (Tabla 1) de esta variable para el grupo control de reflexivos (RC).

Tabla 1
Medias y Desviaciones de las Variables Descriptivas de los Grupos de Seguimiento

Grupo				Edad[a]		CI		Calificación	
	n	*F*	*M*	*M*	*DE*	*M*	*DE*	*M*	*DE*
Autoinstrucción	7	4	3	11.9	0.5	93.1	7.8	8.5	1.0
Habilidad para atender	6	3	3	11.9	0.3	92.5	6.8	8.3	1.1
Impulsivo control	7	3	4	11.9	0.4	93.0	7.7	8.6	1.0
Reflexivo control	10	4	6	11.9	0.5	102.0	8.5	8.9	1.1

[a]Edad cronológica en años.

La Tabla 2 muestra las medias y desviaciones de las latencias y los errores en la tarea MFF durante la postprueba inmediata (postentrenamiento) y en la postprueba de seguimiento (post-30 meses) para los grupos experimentales y controles.

Tabla 2
Medias y Desviaciones de las Variables de la Prueba MFF en el Postentrenamiento (Post) y en el Seguimiento (Post-30 Meses)

Grupo		Latencias				Errores			
		Post		Post-30 meses		Post		Post-30 meses	
	n	*M*	*DS*	*M*	*DS*	*M*	*DS*	*M*	*DS*
Autoinstrucción	7	13.0	17.5	8.1	4.2	10.14	5.7	8.7	2.4
Habilidad para atender	6	13.9	7.6	10.5	4.5	8.8	5.1	7.8	2.6
Impulsivo control	7	3.7	2.0	8.8	2.3	16.1	4.6	9.4	2.7
Reflexivo control	10	16.7	10.9	14.3	5.9	8.4	6.1	6.4	4.6

Se realizaron análisis de los cambios observados en las latencias y los errores MFF del postentrenamiento al post-30 meses (Tabla 1) con la prueba *t* para medidas

repetidas y la prueba de rangos de Wilcoxon. Ambas pruebas indicaron que el incremento de las latencias y el decremento de los errores en el grupo impulsivo control (IC) eran significativos ($p < .05$). En ninguno de los otros grupos se observaron cambios significativos.

Dado que sólo el grupo IC mostró cambios significativos y que no se observaron diferencias significativas (pruebas t con valores de $p > .10$) para las latencias y los errores entre este grupo y los grupos entrenados en autoinstrucciones (AIM) o en la habilidad para atender (HA), se juntaron los tres grupos y se compararon con el grupo RC.

La prueba t para muestras independientes mostró que las latencias de los grupos IC, AIM y HA ($M = 9.1$, $DE = 3.7$), eran menores ($t(28) = 2.99$, $p < .01$) que las del grupo RC y que los errores de este grupo, aunque menores (Tabla 2), no eran estadísticamente diferentes ($t(28) = 1.29$, NS) a los de los otros grupos juntos ($M = 8.5$, $DE = 2.8$).

Estos resultados parecen indicar que los grupos tratados no conservaron los incrementos de latencias de las respuestas aunque sí la exactitud; el hecho de que el grupo IC igualó a los grupos tratados tanto en las latencias como en los errores, en 30 meses, apunta en la misma dirección. En cuanto al aspecto evolutivo, el grupo IC se movió en forma significativa hacia la reflexión (latencias más largas con menor número de errores); sin embargo el grupo RC no se movió hacia la rapidez con exactitud en forma significativa.

Para ver la validez de los resultados obtenidos aplicamos el modelo probabilístico, que establece que aquellos sujetos que cometen 12 o más errores están contestando al azar la tarea MFF, los sujetos que así lo hicieron constituyeron en el pretest el 100% de los impulsivos (grupos control y experimentales) y el 0% de los reflexivos; los porcentajes de sujetos que respondieron al azar en el postentrenamiento y en el post-30 meses, respectivamente, fueron 33% y 0% para el grupo AIM, 44% y 16% (un caso) para el grupo HA, 88% y 14% para el grupo IC y 20% y 10% para el grupo RC. En resumen, podemos decir que en el post-30 meses todos los grupos superaron la tendencia del tiro al azar con variaciones menores en cuanto a las eficiencias individuales.

La clasificación por estilos de los sujetos ($n = 40$) que participaron en el estudio de modificación fue realizada con las medianas del promedio de latencias (7.6 s) y del total de errores (16) de la población de sujetos de la cual fue extraída esa muestra ($N = 87$). Para reclasificar la muestra estudiada en este trabajo, obtuvimos las medianas del promedio de latencias (10.7 s) y del total de errores (9.5) de la población original eliminando los casos que respondieron al azar, de acuerdo al modelo probabilístico.

Dado el problema de la clasificación por estilos con las medianas de latencias y errores, optamos por presentar las latencias individuales de los 30 sujetos estudiados en las tres fases del estudio por grupos (Tabla 3) indicando si respondieron al azar y de lo contrario, si resolvieron la tarea en forma rápida o lenta. Como se puede ver en la Tabla 3, el grupo reflexivo control incluía en preentrenamiento dos sujetos rápidos y ocho lentos, de los cuales cinco (50%) conservaron su lentitud en las siguientes fases. De los dos rápidos uno se conservó como rápido en el post-30 meses y el otro mantuvo sus tiempos cortos, pero pasó a azaroso en las fases de postentrenamiento y post-30

meses. En realidad sólo dos casos previamente reflexivos cambiaron su tiempo de lento a rápido en el post-30 meses.

Tabla 3

Latencias Individuales de los 30 Sujetos en 3 Fases Sucesivas y su Clasificación Como Azarosos, Rápidos o Lentos.

Grupo	Entrenamiento		
	Pre	Post	Post-30 meses
Autoinstrucciones	2.2[a]	5.7[b]	8.5[b]
	1.4[a]	2.7[a]	3.5[b]
	3.5[a]	10.3[b]	13.7[c]
	4.4[a]	5.3[b]	5.3[b]
	6.6[a]	11.7[c]	13.4[c]
	2.9[a]	52.0[c]	8.5[b]
	5.9[a]	3.6[a]	4.1[b]
Habilidad para atender	5.3[a]	13.0[c]	6.8[b]
	7.8[a]	18.5[c]	13.7[c]
	7.6[a]	19.1[c]	10.4[b]
	7.9[a]	22.8[c]	17.8[c]
	4.0[a]	4.0[b]	5.9[b]
	3.6[a]	6.1[a]	8.7[a]
Impulsivo control	4.2[a]	4.5[a]	8.1[b]
	2.4[a]	4.8[b]	7.9[b]
	7.0[a]	5.8[a]	11.7[c]
	3.4[a]	2.7[a]	7.1[b]
	3.2[a]	1.8[a]	10.2[b]
	4.5[a]	2.6[a]	11.1[a]
	7.7[a]	3.4[a]	5.4[b]
Reflexivo control	12.3[c]	9.3[b]	14.9[c]
	11.8[c]	14.5[c]	9.0[b]
	16.6[c]	18.1[c]	7.7[b]
	32.2[c]	41.8[c]	20.6[c]
	12.2[c]	16.8[a]	22.6[c]
	26.4[c]	23.7[c]	21.2[c]
	9.0[b]	5.1[a]	6.7[a]
	14.6[c]	21.2[c]	13.7[c]
	14.1[c]	11.7[c]	16.9[c]
	9.9[b]	4.1[a]	9.8[b]

Nota. Las latencias son en segundos. a = azarosos, b = rápidos, c = lentos.

En los grupos impulsivos se puede observar que en la fase de postentrenamiento, los grupos entrenados dejaron de responder al azar en un 77% de los sujetos, mientras que

en el grupo IC sólo el 14% (un caso) no respondió al azar. Sin embargo, 30 meses después sólo el 10% (un caso del IC y uno del HA) de todos los sujetos impulsivos continuaba respondiendo al azar y los tiempos de los que no respondieron al azar se clasificaron con un mayor porcentaje (72%) como rápidos que como lentos (28%). Finalmente, si consideramos únicamente a los grupos que tuvieron entrenamiento, 67% quedaron clasificados como rápidos y 33% como lentos en el post-30 meses.

DISCUSIÓN

El primer objetivo de este estudio fue buscar evidencia de la duración de los efectos de modificación del estilo impulsivo logrado por dos métodos radicalmente diferentes: el método de autoinstrucciones con modelamiento en reflexión y el de incremento de la habilidad para atender (Solís-Cámara, 1985). Del análisis de los resultados de la Tabla 2 parecería poderse concluir que las latencias entre el postentrenamiento y el post-30 meses se conservaron globalmente para los grupos entrenados, de la misma manera que para el grupo reflexivo control; mientras que el grupo impulsivo control sí mostró incremento significativo en sus latencias hasta igualarse con las de los grupos entrenados, pareciendo así haberse hecho más reflexivo (latencias más largas y menos errores); sin embargo, los tres grupos impulsivos mostraron diferencia significativa con las latencias del grupo reflexivo control, con lo que se contradice el supuesto desplazamiento de los impulsivos hacia la reflexión.

El segundo objetivo de este estudio fue el de ver si los grupos impulsivo control y reflexivo control conservaban su estilo o si el primero se movía a reflexivo y el segundo a rápido-exacto de acuerdo a lo informado en la literatura (Salkind et al., 1978). De nuevo el análisis de los resultados presentados en la Tabla 2 parece indicar que el grupo impulsivo mantuvo su estilo, aunque aumentó sus latencias y disminuyó sus errores, porque conservó latencias significativamente más cortas que las del grupo control reflexivo; sin embargo, ambos grupos no mostraron diferencias significativas en cuanto a la variable total de errores, lo que contradice las diferencias esperadas para estos estilos, según lo informado en la literatura (Messer, 1976).

Contradicciones como las arriba señaladas no son poco comunes en la literatura de los estilos de reflexión–impulsividad y llevaron a los autores a proponer que no se consideren estilos cognoscitivos aquellos casos en que los sujetos al contestar tienen una probabilidad significativa de haberlo hecho al azar (total de errores igual o mayor a 12 errores; Solís-Cámara & Solís-Cámara, 1987). Analizados retrospectivamente los datos del estudio de modificación (Solís-Cámara, 1985), con este criterio, el 100% de los impulsivos no mostraban un tiempo conceptual impulsivo por lo que no los consideramos como impulsivos cognoscitivos, sino que eran sujetos que contestaron la tarea respondiendo al azar. No así el grupo reflexivo control, en el que los sujetos sí podían ser clasificados por estilos. A este respecto, la clasificación por estilos según las medianas queda sujeta a las características propias de cada grupo estudiado, particu-

larmente en cuanto al porcentaje de niños que contestan la tarea respondiendo al azar y que hemos mostrado que varía en forma proporcional inversa a la edad de los niños (Solís-Cámara & Solís-Cámara, 1987). Es probable que la mayor impulsividad (latencias menores con muchos errores) encontrada en preescolares y en niños de nivel socioeconómico bajo y rural sea en realidad reflejo de respuestas al azar (Messer, 1976; Solís-Cámara & Gómez, 1985; Solís-Cámara & Solís-Cámara, 1986b).

Por lo anterior, los autores han propuesto que sean usadas las medianas únicamente de aquellos sujetos que sí usan el proceso cognoscitivo de evaluación de hipótesis alternativas (Kagan, Lapidus & Moore, 1978) en la tarea MFF (menos de 12 errores). Clasificados con este criterio el grupo reflexivo control en el preentrenamiento (Solís-Cámara, 1985) resultó preponderantemente lento-inexacto e incluyendo a dos *impulsivos cognoscitivos*. Esta forma de clasificar a los sujetos por estilos no supera, sin embargo, el problema de que al separar a los sujetos por las medianas, un sujeto puede cambiar de estilo por sólo una diferencia de un error o una fracción de segundo. Por ello decidimos minimizar los estilos (en este trabajo) y clasificarlos como lentos o rápidos, lo que está de acuerdo con el postulado original de Kagan (Kagan et al., 1964). Creemos que el análisis de los casos individuales (Tabla 3) más los análisis globales pueden dar mejor sentido a las inferencias estadísticas en estudios de seguimiento.

A la luz del modelo probabilístico, la interpretación retrospectiva del estudio de modificación del estilo impulsivo nos lleva a las siguientes conclusiones.

1. No se modificó el *estilo* impulsivo o la impulsividad cognoscitiva sino la tendencia de responder a la tarea al azar. Esto explica el que un método no cognoscitivo como el de desarrollar la capacidad para atender haya logrado resultados semejantes a los del método cognoscitivo de autoinstrucciones. El grupo impulsivo control siguió contestando la tarea al azar en el postentrenamiento.

2. En el postentrenamiento los grupos entrenados presentaron 6 (54%) casos de lentos, de los 11 casos que definieron su estilo. En el post-30 meses sólo 4 (33%) de los 12 que definieron su estilo fueron casos lentos, la mayoría mostró una tendencia a regresar al uso de tiempos cortos. En el grupo impulsivo control de 6 casos que definieron su estilo en el post-30 meses, 5 fueron rápidos (83%). Lo que parece indicar que el efecto del entrenamiento sobre las latencias se desvaneció gradualmente a largo plazo y los tiempos cortos fueron la característica sobresaliente de todos los grupos impulsivos del estudio de seguimiento.

3. En cuanto al grupo reflexivo control, se ve que éste resultó contaminado con la presencia de dos casos rápidos, que los tiempos cortos de estos dos casos favorecieron que en el postentrenamiento se perdieran las diferencias por latencias entre los impulsivos entrenados y este grupo control. En el post-30 meses, estos dos casos contribuyeron al promedio del total de errores con 16 y con 11 errores cuando el valor máximo de los restantes sujetos fue de 8 y, que si quitamos estos dos casos, el promedio de latencias se alarga ($M = 15.85$, $DE =$

5.6) y el de errores disminuye ($M = 4.62$, $DE = 2.8$) con lo que las latencias se mantienen diferentes a las de los 3 grupos impulsivos ($t(26) = 3.76$, $p < .01$) y las diferencias en errores se hacen significativas ($t(26) = 2.96$, $p < .01$). Globalmente y de acuerdo a las mismas medianas del grupo que no respondió al azar, este grupo queda mejor caracterizado como reflexivo a pesar de la presencia de dos casos que acortaron sus tiempos en el post-30 meses.

En resumen, los entrenamientos sirvieron para modificar la respuesta al azar de los sujetos y los cambios en latencia logrados se desvanecieron con el paso del tiempo. El grupo impulsivo control se movió a impulsivo cognoscitivo igualando sus tiempos al de los grupos entrenados mientras el grupo reflexivo control definió mejor su estilo. Una vez definidos los estilos, los sujetos con tiempos cortos tendieron a mantener sus tiempos cortos y los de tiempos largos sus tiempos largos.

Este es un estudio preliminar en el que se proporciona evidencia de la importancia de descartar en los estudios de reflexión–impulsividad a los sujetos que responden al azar a la tarea MFF y que llevan a interpretaciones inexactas de los resultados. Consideramos que estudios de modificación del estilo impulsivo como el realizado por los autores (Solís-Cámara, 1985) y por otros (Meichembaum & Goodman, 1971) deben ser repetidos en impulsivos cognoscitivos, ya que nuestros resultados no apoyan la modificación duradera del estilo. El estilo cognoscitivo parece ser una disposición del niño, no una capacidad. Asimismo, en los estudios acerca del tiempo conceptual debe tomarse en cuenta que las medianas de latencia no pueden separar en forma pura a los rápidos de los lentos ya que ambos grupos tienen la capacidad de moverse en la elección de tiempos más o menos cortos o más o menos largos, es decir, cualquier mediana va a excluir a algunos rápidos o lentos y a incluir por el contrario a otros lentos o rápidos, según la población elegida.

Finalmente, establecer con mayor claridad las propiedades psicométricas de la prueba MFF es de suma importancia, pues parece poco razonable el que la prueba tal como está diseñada sea usada como índice de una dimensión cognoscitiva en niños de 6 a 12 años de edad (Kagan et al., 1978), un lapso extremadamente largo. Por otra parte, una última consideración con respecto al llamado estilo impulsivo, según MFF, es que es aceptado o usado indistintamente con el concepto tradicional de impulsividad clínica o desórdenes del impulso, lo cual es definido: acto no premeditado o respuestas precipitadas, "sin pensar" (¿al azar?; ej., ver Schaefer & Millman, 1980); tal vez el grupo de niños que responden al azar en la tarea MFF corresponda a esa entidad clínica vagamente definida, aunque popular. Sin embargo, la definición de estilos cognoscitivos de R–I de Kagan no indica que ésa haya sido la intención al establecer tales estilos (Kagan et al., 1964; Kagan & Messer, 1975) ni al reconfirmar las situaciones en las que la dimensión es válida (Kagan et al., 1978).

En conclusión, la mejor definición de los estilos llevaría a evitar considerar a los individuos reflexivos en MFF como el paradigma del desempeño de los niños normales o maduros y a evitar la modificación del tiempo del estilo impulsivo, ya que la eficiencia de cada estilo dependería de las demandas situacionales de cada tarea en particular y el estilo no mostraría estar asociado a un desempeño ineficiente o eficiente generalizado.

Por otra parte, los sujetos que responden al azar se beneficiarían con los procedimientos de entrenamiento para corregir su ineficiencia en tareas de selección múltiple y probablemente en otras situaciones con alta incertidumbre de la respuesta correcta. Dado lo pequeño de la muestra estudiada, no fue posible realizar estadísticas que parecen necesarias como análisis de correlación y de regresión. Lo interesante de los resultados hace necesario confirmarlos con estudios longitudinales con niños definidos como rápidos (impulsivos cognoscitivos), lentos (reflexivos cognoscitivos) o como individuos que responden al azar, y conducir estudios transversales descriptivos que permitan definir más claramente la dimensión de reflexión–impulsividad en sujetos que no responden al azar, así como determinar las deficiencias del grupo que responde azarosamente.

REFERENCIAS

Ancillotti, J.P. (1985). Dimension réflexive–impulsive de la personnalité et fonctionnement cognitif. *Cahiers de Psychologie Cognitive, 5*, 667–686.

Block, J., Block, J.H., & Harrington, D.M. (1974). Some misgivings about the Matching Familiar Figures Test as a measure of reflection–impulsivity. *Developmental Psychology, 10*, 611–632.

Block, J., Block, J.H., & Harrington, D.M. (1975). Comment on the Kagan–Messer reply. *Developmental Psychology, 11*, 249–252.

Gjerde, P.F., Block, J., & Block, J.H. (1985). Longitudinal consistency of the Matching Familiar Figures Test from early childhood to preadolescence. *Developmental Psychology, 21*, 262–271.

The Goodenough Test (1982). *Standardized version for school-age children in Guadalajara City*. Public Education State Board. Departamento de Educación Pública del Estado. Servicio de Psicopedagogía. Escuela Normal de Jalisco. Guadalajara, México.

Kagan, J. (1971). *Understanding children: Behavior, motives and thought*. New York: Harcourt Brace Jovanovich.

Kagan, J., Lapidus, D.R., & Moore, M. (1978). Infant antecedents of cognitive functioning: A longitudinal study: *Child Development, 49*. 1005–1023.

Kagan, J., & Messer, S.B. (1975). A reply to: "Some misgivings about the Matching Familiar Figures Test as a measure of reflection–impulsivity." *Developmental Psychology, 11*, 244–248.

Kagan, J., Rossman, B.L., Day, D., Albert, J., & Phillips, S. (1964). Information processing in the child: Significance of analitic and reflective attitudes. *Psychological Monographs, 78*, (1, Whole No. 578).

Kogan, N. (1976). *Cognitive styles in infancy and early childhood.* New Jersey: Erlbaum.

Meichembaum, D. & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77,* 115–126.

Messer, S.B. (1976). Reflection–impulsivity: A review. *Psychological Bulletin, 83,* 1026–1052.

Salkind, N.J., Kojima, H., & Zelniker, T. (1978). Cognitive tempo in American, Japanese, and Israeli children. *Child Development, 49,* 1024–1027.

Salkind, N.J., & Wright, J.C. (1977). The development of reflection–impulsivity and cognitive efficiency. *Human Development, 20,* 377–387.

Schaefer, C.E., & Millman, H.L. (1980). *Therapies for children.* San Francisco, CA: Jossey-Bass.

Solís-Cámara R., P. (1985). Efectos del entrenamiento en la habilidad para atender, versus el uso de autoinstrucciones en la modificación de la impulsividad. *Revista Latinoamericana de Psicología, 17,* 205–226.

Solís-Cámara R., P., & Gómez M., L. (1985). Children's human figure drawings and impulsive style at two levels of socioeconomic status. *Perceptual and Motor Skills, 61,* 1039–1042.

Solís-Cámara R., P., & Solís-Cámara V., P. (1986a). Algunas evidencias a favor de la contaminación motriz del tiempo conceptual. *Salud Mental, 9,* 72–79.

Solís-Cámara R., P. & Solís-Cámara V., P. (1986b). El desarrollo de la eficiencia de escolares de dos niveles socioeconómicos en una tarea de estilos. *Salud Mental, 9,* 61–67.

Solís-Cámara R., P., & Solís-Cámara V., P. (1987). Is the Matching Familiar Figures Test a measure of cognitive style? A warning for users. *Perceptual and Motor Skills, 64,* 59–74.

Solís-Cámara R., P., Troyo S., R., & Solís-Cámara V., P. (1983). Relación entre los estilos de reflexión–impulsividad y el desempeño de niños en una tarea de discriminación visual sucesiva. *Salud Mental, 6,* 24–32.

Recibido en octubre 3 de 1986
Revisión recibida en junio 16 de 1987

AUTORES

PEDRO SOLÍS-CÁMARA RESÉNDIZ. *Afiliación institucional:* Instituto Mexicano del Seguro Social, Guadalajara, Jalisco, México. *Título:* Investigador (Sección de Psicología). *Diplomas:* Licenciado en Psicología, 1977, Universidad Autónoma de San Luis Potosí, México; Master of Arts, 1982, Ohio State University, U.S.A. *Especializaciones:* Diagnóstico y tratamiento de incapacidades en el desarrollo infantil; manejo del comportamiento infantil. *Dirección de correo:* Apartado Postal No. 2-322, Guadalajara, Jalisco, México.

PEDRO SOLÍS-CÁMARA VALENCIA. *Afiliación institucional:* Instituto Mexicano del Seguro Social, Guadalajara, Jalisco, México. *Título:* Investigador (Sección de Psicología). *Diplomas:* Médico Cirujano, 1959, Universidad Nacional Autónoma de México. *Especialización:* Fisiología humana. *Dirección de correo:* Apartado Postal No. 2-322, Guadalajara, Jalisco, México.

EXTENDED SUMMARY

Long-term (30 Months) Effects of Attempts to Modify Impulsivity: Their Relevance to an Understanding of Style

PEDRO SOLÍS-CÁMARA R.

and

PEDRO SOLÍS-CÁMARA V.

In a 1985 study (Solís-Cámara) the authors compared the efficiency of two different procedures for modifying the behavior of impulsive children in a visual discrimination task: (a) the use of self-instructions with reflective modeling (Meichembaum & Goodman, 1971) and (b) training in the ability to pay attention. Both procedures had proved to be effective strategies in altering the impulsivity style. This finding did not agree with Kagan and Messer's (1975) conceptualization of conceptual tempo nor with misgivings (Block, Block, & Harrington, 1974) concerning the reflection–impulsivity dimension and its measure, the Matching Familiar Figures Test (MFFT). The authors also noted that none of the explanations of the simultaneous increase of response latencies and the decrease of total errors with age, *based on group performances on the MFFT*, has proved to be empirically reliable. Seeking an explanation of the contradictions found in several studies, we recently proposed a probabilistic model (Solís-Cámara & Solís-Cámara, 1987). This model shows that some children are, in a statistically significant manner, giving random responses (12 or more total errors) on the MFFT and thus are contaminating the performances of other children. The present study tests this model.

A follow-up investigation was undertaken of the children who had participated in the study of modification of the impulsivity style. The objective was to know whether impulsive children, treated to become reflectives, kept their altered styles after 30 months without intervention. We also wanted to know whether nontreated children retained their previous styles or showed a developmental trend with age (30 months).

Twenty impulsive and 10 reflective subjects participated in this study. They were drawn from an initial sample of 38 children selected on the basis of their MFFT scores (Solís-Cámara, 1985). Subjects were in sixth grade at the time of this study. Each group was composed as follows: a group ($n = 7$) originally trained in self-instructions with reflective modeling; a group ($n = 6$) trained to pay attention; a group of impulsive children to function as a control ($n = 7$), and a group of reflectives also to function as a control ($n = 10$). Of this sample 14 were girls and 16 boys. Mean age was 11 years and 11 months; mean IQ was 95.9 ($SD = 7.8$). The school's mean mark was 8.6 ($SD = 1.0$, on a zero to ten scale). Children came from urban sections of Guadalajara City, México, and were drawn from middle to low income families.

Results indicated that impulsives treated, by either one of the two different procedures, to become reflectives did not keep their longer latencies but only their accuracy (fewer errors). The expected developmental trend was not found. When the probabilistic model was used a different scene emerged. In retrospect, the subjects who randomly responded to the MFFT during its first administration (pretest) were the 100% impulsives and the 0% reflectives. In the posttraining the impulsives of the control group kept answering randomly on the test (100%) while just a few (23%) of the trained impulsives did so. Most subjects in each group did not give random responses 30 months later.

Further analyses in which we used the probabilistic model allowed us to separate random respondents from fast and slow respondents. Results indicated that when subjects are classified on the basis of the reflective-impulsive (R–I) style dimension, and random respondents are not included in the analyses, most children retain their slow or fast conceptual tempo. These preliminary results support the need for additional studies of the R–I dimension using the probabilistic model.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp.56-71

Efeitos de um Programa de Treinamento de Criatividade Para Professores do Ensino de Primeiro Grau nas Habilidades de Pensamento Criativo do Aluno

EUNICE MARIA LIMA SORIANO DE ALENCAR
DENISE DE SOUZA FLEITH
LOURDES A. SHIMABUKURO
e
MÁRCIA DE ANDRADE NOBRE

Universidade de Brasília
BRASIL

Investigaram-se no estudo os efeitos de um programa de treinamento de professores de 1º grau para a estimulação das habilidades de pensamento criativo no nível de criatividade de seus alunos e em interesses e características associadas à criatividade. A amostra foi constituída por 42 professores. Deste total, 23 participaram do programa de criatividade e 19 constituíram o grupo de controle. Pré e pós-testes de criatividade escolhidos dentre os Testes de Pensamento Criativo de Torrance foram aplicados em uma amostra de 140 alunos dos professores de ambos os grupos, escolhidos aleatoriamente. Estes alunos responderam ainda a um inventário, informando a respeito de interesses e características que supostamente fundamentam o comportamento criativo. Os resultados indicaram que, embora em 10 das 12 medidas de criatividade os alunos dos professores que participaram do treinamento apresentassem um desempenho superior, tais diferenças não foram significativas. Diferenças significativas foram, porém, observadas em 6 dentre as 12 medidas de criatividade a favor dos alunos do sexo feminino, independentemente de seus professores terem participado ou não do programa de treinamento de criatividade. Observou-se que uma alta porcentagem de alunos responderam afirmativamente a itens relativos a características que se associam a criatividade, como alta imaginação e curiosidade.

Trabalho apresentado na 37ª Reunião Anual da Sociedade Brasileira Para o Progresso da Ciência, Curitiba, julho de 1986. O mesmo foi parcialmente financiado pelo CNPq (Processo nº 407672/84 ED).

Copyright © 1987 by E.M.L. Soriano de Alencar, D. de Souza Fleith, L.A. Shimabukuro, and M. de Andrade Nobre.
All rights reserved.

A necessidade e importância de se criarem condições favoráveis ao desenvolvimento da criatividade do aluno têm sido apontadas por pesquisadores de diferentes países. Dentre as razões que poderiam ser lembradas para justificar tal necessidade salientam-se: (a) Problemas e desafios. A criança de hoje vai lidar com inúmeros problemas e desafios durante toda a sua vida; possivelmente, ela enfrentará individualmente mais problemas do que é possível ensiná-la a resolver durante todos os seus anos de escola. (b) Satisfação e prazer. Como salientado pela psicologia humanística, a atualização do potencial é uma experiência gratificante para o indivíduo. Na medida em que a aprendizagem se dá de uma forma criativa, isso favorece o crescimento do aluno e gera prazer, motivando-o para novos avanços (Treffinger, 1979).

Um dos fatores que também contribuiu para chamar a atenção para a necessidade de melhores condições ao desenvolvimento da criatividade foi o movimento da potencialidade humana, o qual aponta para o fato de que o ser humano tem feito uso de apenas uma parcela muito limitada de seu potencial criador. Pesquisas na área de *biofeedback*, da parapsicologia, de estados alterados da consciência atestam este imenso potencial que, na maioria das vezes, permanece inibido, adormecido ou bloqueado (Houston, 1977).

No Brasil, de modo geral, os professores do ensino de primeiro grau não oferecem as condições mais adequadas para o desenvolvimento da criatividade do aluno. Observa-se que várias são as idéias errôneas a respeito de criatividade que predominam entre muitos educadores. Uma dessas idéias é a de que a criatividade ocorreria apenas em atividades artísticas, cabendo ao professor de artes o seu desenvolvimento. Observa-se ainda que o aluno mais criativo não é reconhecido na escola e nem tampouco tem recebido uma atenção maior por parte de seus professores. Dado o programa extenso da escola brasileira, o curto período de tempo que o aluno permanece na escola e a falta de treinamento dos professores, tende-se a enfatizar a memorização de informações, com pouca ênfase na aplicação do conhecimento e no uso do pensamento divergente.

Por outro lado, nota-se que para favorecer o desenvolvimento das habilidades criativas do aluno, são necessárias tanto condições ambientais favoráveis como o domínio de técnicas adequadas. Com relação a este aspecto, observa-se que muitos foram os métodos e programas desenvolvidos para serem utilizados tanto por adultos como por crianças. Os mais conhecidos são o método *brainstorming* desenvolvido por Osborn (1963) e o método da sinética, desenvolvido por Gordon (1971). Para alunos do ensino de 1º grau salientam-se o Programa de Pensamento Produtivo desenvolvido por Covington, Crutchfield e Davis (1966) e o Programa de Pensamento Criativo de Purdue (Feldhusen, Treffinger e Bahlke, 1970).

Este último foi utilizado por nós em um estudo com alunos de 4ª e 5ª séries de escolas públicas e particulares de Brasília, DF, apresentando as crianças, ao término do programa, um desempenho em testes de criatividade significativamente superior àquele apresentado pelos sujeitos que constituíram o grupo de controle (Alencar, 1974a).

No Brasil, outras pesquisas sobre treinamento de criatividade foram desenvolvidas por Marin (1976) e por Wechsler (no prelo)[1]. A pesquisa de Marin teve como objetivo

[1] Outros estudos foram também desenvolvidos, como o de Guhur (1976) e Maia (1975). Estes foram, porém, estudos de curta duração, utilizando-se ainda poucas medidas de criatividade.

investigar o efeito de um programa de atividades artísticas na criatividade não-verbal de 37 alunos que cursavam a 1ª série do segundo grau em uma escola pública. Dois testes de criatividade da bateria Torrance foram aplicados tanto antes do início do programa como após o seu término a estes sujeitos e a outros 37 sujeitos que constituíram o grupo de controle. Marin observou que os alunos que participaram das atividades de educação artística apresentaram, ao final do programa, ganhos significativos em medidas de fluência, flexibilidade e originalidade, demonstrando assim que o programa de educação artística influiu positivamente sobre a criatividade não-verbal dos estudantes.

Também Wechsler (no prelo) investigou os efeitos do treinamento de criatividade em uma amostra de crianças bem-dotadas e outra de crianças que não se destacavam por um desempenho ou habilidades superiores. Para a seleção destes sujeitos, foram utilizados testes de criatividade de natureza verbal e não verbal da bateria Torrance, testes de inteligência (Raven e Desenho da Figura Humana) e questionários respondidos tanto pelas professoras como pelas crianças. O treinamento de criatividade consistiu em 23 sessões com duração de duas horas cada, as quais ocorriam duas vezes por semana. Durante estas sessões, exercícios de pensamento divergente e outros que visavam o desenvolvimento integral do aluno foram ministrados. Ganhos em medidas de criatividade e no desempenho escolar foram observados entre todos os sujeitos que participaram do programa. Observou-se que nas variáveis de fluência e originalidade de natureza não-verbal, fluência verbal e motivação, os ganhos das crianças bem-dotadas foram inferiores àqueles apresentados pelas demais crianças. Constatou-se ainda que os efeitos do treinamento criativo foram mais acentuados nas variáveis de originalidade não-verbal, motivação, participação e rendimento.

Os estudos realizados com programas de criatividade aplicados em escolas do ensino de 1º grau têm indicado, pois, ganhos substanciais em habilidades de pensamento criativo por parte dos alunos. Entretanto, é nosso ponto de vista que, para a manutenção desses ganhos, é necessário que estes programas não sejam vistos como um corpo estranho em sala de aula, mas antes que haja uma assimilação por parte dos professores dos princípios básicos que os fundamentam, passando os professores a fazer uso dos mesmos e incentivando-os entre os seus alunos. Caso contrário, todos os ganhos observados através de sua utilização declinarão rapidamente com o tempo, até desaparecer completamente.

Os dados obtidos por nós em várias pesquisas na área de criatividade (Alencar, 1974b e c; 1975b, 1984, 1985) sugerindo algumas delas a influência do professor na criatividade do aluno e outras a presença de um ambiente pouco favorável ao desenvolvimento da criatividade do aluno em sala de aula, levaram-nos a concluir sobre a necessidade de se desenvolver na escola um maior número de atividades com o objetivo de se treinar o aluno na produção de novas idéias e na produção criativa em diferentes campos de estudo. Tal treinamento é necessário para um melhor desenvolvimento das habilidades cognitivas de nossos alunos.

Neste sentido, um treinamento de professores no uso de pensamento criativo seria especialmente útil, uma vez que os nossos professores têm pouco conhecimento ou compreensão dos processos de pensamento criativo e de como ensiná-lo. Interessados

neste aspecto, desenvolvemos o presente estudo, cujos objetivos foram: (a) investigar os efeitos de um programa de criatividade para professores do ensino de 1° grau no nível de criatividade de seus alunos; (b) investigar as diferenças entre alunos do sexo masculino e feminino em suas habilidades de pensamento criativo; e (c) investigar as interações entre as variáveis programa de criatividade e sexo do aluno.

MÉTODO

Sujeitos

A amostra foi constituída por 42 professores do ensino de 1° grau de escolas públicas da cidade-satélite do Gama, Distrito Federal; 23 professores participaram do programa de criatividade e 19 constituíram o grupo de controle. Todos os professores que participaram do estudo eram do sexo feminino e lecionavam em turmas de 4ª série de escolas públicas da cidade-satélite do Gama, da zona rural adjacente a esta cidade ou trabalhavam no Complexo Escolar desta cidade, que se situa cerca de 40 km de Brasília, DF. Dos 23 professores que constituíram o grupo experimental, 18 eram casados e 5 solteiros. A sua idade média era 35,2 anos e tinham em média 10,6 anos de experiência de ensino; 14 tinham o curso de magistério e 9 tinham ainda o curso de pedagogia e/ou outros cursos superiores. O número médio de alunos por turma era 40,1.

A escolha da cidade do Gama para o desenvolvimento do projeto foi feita pela Diretoria Geral de Pedagogia da Fundação Educacional do Distrito Federal, que utilizou como critério o fato de serem os professores desta cidade aqueles com menos oportunidades de participarem em cursos ou outras atividades promovidas pela Fundação Educacional. A escolha dos professores que iriam participar do treinamento de criatividade foi feita pela Diretoria do Complexo Escolar do Gama. Apesar de ter sido enfatizado que todos os professores deveriam lecionar na 4ª série, dos 23 professores que participaram do treinamento, 3 eram dinamizadores[2] e 4 trabalhavam no Complexo Escolar em atividades administrativas. Uma das professoras que participou do treinamento, lecionou apenas poucas semanas no início do semestre em uma turma de 4ª série. Desta forma, somente 15 professores que participaram do treinamento lecionavam em turmas de 4ª séries durante todo o semestre em que o programa de criatividade foi levado a efeito. Os professores do grupo de controle foram escolhidos dentre aqueles que lecionavam em turmas de 4ª séries das mesmas escolas onde lecionavam os professores do grupo experimental.

De cada uma das turmas de 4ª série dos 34 professores (15 do grupo experimental e 19 do grupo de controle), foram selecionados aleatoriamente 5 alunos para responder ao pré-teste (170 alunos no total). No dia da aplicação do pós-teste (no final de novembro), 30 alunos (16 do grupo experimental e 14 do grupo de controle) deste total não estavam presentes na escola, o que reduziu para 140 alunos (59 do grupo

[2] Professores dinamizadores são aqueles que não têm uma turma fixa, permanecendo um dia da semana em cada sala de aula, desenvolvendo atividades ligadas à educação artística, recreação, ou substituindo professores ausentes. Salienta-se que só tomamos conhecimento da inclusão destes professores após a aplicação do pré-teste, no primeiro encontro de criatividade.

experimental e 81 do grupo de controle) os participantes do estudo. Deste total, 65 eram do sexo masculino e 75 do sexo feminino e sua idade média era 11,1 anos.

O Programa de Criatividade

Os professores do grupo experimental participaram de dez encontros, com duração de duas horas e trinta minutos cada. Estes encontros ocorreram uma vez por semana em uma sala de aula de uma escola onde funcionava o Complexo Escolar do Gama.

Os seguintes tópicos foram aí abordados: Idéias errôneas sobre criatividade; barreiras ao desenvolvimento da criatividade; influência do ambiente familiar sobre a criatividade; traços de personalidade associados à criatividade; criatividade como conjunto de habilidades cognitivas; fatores que constituem barreiras ao desenvolvimento da criatividade na escola; a identificação dos indivíduos criativos; programas de treinamento de criatividade; elaboração de miniprojetos específicos para o desenvolvimento de habilidades de pensamento criativo em sala de aula.

Atividades práticas foram desenvolvidas durante os encontros com o objetivo de treinar o professor: (a) na produção de múltiplas idéias para várias tarefas cognitivas; (b) na criação de idéias originais e inovadoras; (c) na avaliação de idéias ou soluções alternativas para situações de problemas; (d) na elaboração de projetos criativos em pequenos grupos.

Durante cada encontro, foi ainda distribuído um pequeno texto para ser lido e discutido pelos professores. Estes foram elaborados pela pesquisadora. Constituíram uma versão simplificada dos principais tópicos do programa extraído do texto *Psicologia da criatividade* de Alencar (1986). Ao final do programa, todos os professores receberam un certificado de freqüência ao programa de criatividade.

Avaliação da Criatividade do Aluno

Tanto antes como imediatamente após o término do programa de criatividade, dois testes de natureza verbal e dois testes de natureza figurativa, foram aplicados aos 140 alunos dos professores que constituíram o grupo experimental e o grupo de controle. Os seguintes testes, escolhidos dentre os Testes de Pensamento Criativo de Torrance (Torrance, 1966), foram utilizados: (a) Círculos; (b) Usos Inusuais; (c) Complementação de Figuras e (d) Aperfeiçoamento de Produto.

No teste Círculos, o aluno recebe uma série de círculos e é instruído a fazer o maior número de objetos ou de desenhos diferentes com eles. Em Usos Inusuais, o aluno deve dar o maior número possível de usos novos e inusuais para latas. Em Complementação de Figuras, o aluno recebe 10 figuras incompletas e é instruído a completar cada uma delas. Em Aperfeiçoamento de Produto, pede-se a ele a produção de maneiras inteligentes e incomuns de aperfeiçoar um pequeno brinquedo. As crianças são encorajadas em todos os testes a apresentar idéias ou respostas que não sejam dadas

pelos seus colegas.

Cada um dos quatro testes foi avaliado nas categorias de fluência, flexibilidade e originalidade, usando-se os procedimentos recomendados por Torrance (1966). Fluência é o número total de respostas relevantes, relevância sendo definida em termos dos requisitos das tarefas como apresentadas nas instruções. Flexibilidade é o número total de diferentes categorias em que as respostas dos sujeitos podem ser classificadas. Originalidade se baseia na raridade estatística das idéias expressas. A raridade estatística é a freqüência de ocorrência das respostas na população de respostas. Escores de 0, 1 e 2 foram dados dependendo da originalidade da resposta. Os escores de originalidade foram baseados na raridade estatística das respostas na amostra estudada. Uma amostra randômica de 100 sujeitos foi usada com este propósito. Foi escolhido este teste, pelo fato de virmos utilizando o mesmo em nossas pesquisas na área de criatividade.

Observa-se, entretanto, que embora Torrance (1972) apresente dados que indiquem níveis satisfatórios de validade de sua bateria, alguns autores têm levantado alguns problemas com relação aos testes de criatividade (neste sentido, veja Petrosko, 1978). Por esta razão, além de responder aos testes de criatividade, os alunos selecionados responderam também a um inventário onde informaram a respeito de interesses e características que supostamente fundamentam o comportamento criativo. Na elaboração deste inventário, aproveitaram-se alguns itens sugeridos por Davis (1975) e outros foram acrescentados. Para ilustração, apresentamos a seguir alguns itens incluídos neste instrumento: Eu sou muito curioso; eu tenho muita imaginação; eu nem sempre penso da mesma forma que os meus colegas; na minha opinião, quando se desenha o sol, ele deve ser sempre pintado de amarelo (relação inversa com maior criatividade).

RESULTADOS

Foi o objetivo principal do estudo investigar o efeito do programa de treinamento de criatividade no nível de criatividade dos alunos. Na análise dos dados relativos a este aspecto, utilizaram-se análises de covariância, com dois fatores. O primeiro — condição experimental— envolveu dois níveis: programa de treinamento de criatividade e controle. O segundo fator foi sexo, com dois níveis. As variáveis dependentes foram as doze medidas nos testes de criatividade, ou seja, os escores de fluência, flexibilidade e originalidade em dois testes de natureza figurativa e os escores de fluência, flexibilidade e originalidade em dois testes de natureza verbal. Os escores obtidos no pré-teste serviram como covariantes para os escores do pós-teste.

Na Tabela 1, são apresentadas a média global, a do grupo experimental e a do grupo de controle (para simplificar, algumas vezes utilizaremos os termos grupo experimental e de controle para se referirem aos alunos dos professores que constituíram estes grupos) nas doze medidas de criatividade. Como pode ser observado nesta Tabela, em

10 das 12 medidas o grupo experimental obteve médias superiores àquelas obtidas pelo grupo de controle. Entretanto, através da anălise de covariância, observou-se que tais diferenças não foram significativas.

Tabela 1

Médias Obtidas Pelo Grupo Total, Pelo Experimental e Pelo Grupo de Controle nas Medidas de Fluência, Flexibilidade e Originalidade nos Quatro Testes de Criatividade (Pós-Teste)

Testes	Medidas	Exper.	Contr.	Total
Círculos	Fluência	10,93	10,12	10,46
(figurativo)	Flexibilidade	8,71	8,38	8,52
	Originalidade	7,34	6,37	6,78
Usos Inusuais	Fluência	17,92	18,49	18,25
(verbal)	Flexibilidade	5,92	5,65	5,76
	Originalidade	5,63	6,09	5,89
Complem. Figuras	Fluência	9,32	9,16	9,23
(figurativo)	Flexibilidade	8,29	8,00	8,12
	Originalidade	9,64	9,19	9,38
Aperf. Produto	Fluência	11,76	10,58	11,08
(verbal)	Flexibilidade	4,56	4,09	4,29
	Originalidade	8,07	7,35	7,65

As médias obtidas pelos sujeitos do sexo masculino e feminino são apresentadas na Tabela 2. Nota-se pela tabela, que em ll dentre as 12 medidas de criatividade os alunos do sexo feminino obtiveram médias superiores às do sexo masculino (seis destas diferenças foram estatisticamente significativas).

Tabela 2

Médias Obtidas Pelos Sujeitos do Sexo Masculino e Feminino nas Medidas de Fluência, Flexibilidade e Originalidade em Quatro Testes de Criatividade (Pós-Teste)

Testes	Medidas	Masculino	Feminino
Círculos	Fluência	9,11	11,64
(figurativo)	Flexibilidade	7,58	9,33
	Originalidade	5,66	7,75
Usos Inusuais	Fluência	15,98	20,21
(verbal)	Flexibilidade	5,62	5,89
	Originalidade	5,32	6,39
Compl. Figuras	Fluência	8,98	9,44
(figurativo)	Flexibilidade	8,03	8,20
	Originalidade	9,57	9,21

Médias Obtidas Pelos Sujeitos do Sexo Masculino e Feminino nas Medidas de Fluência, Flexibilidade e Originalidade em Quatro Testes de Criatividade (Pós-Teste) (Cont.)

Testes	Medidas	Masculino	Feminino
Aperf. Produto (verbal)	Fluência	9,31	12,61
	Flexibilidade	4,17	4,39
	Originalidade	6,57	8,59

Através da análise de covariância, observou-se que o valor *F* obtido para o efeito principal da condição experimental não foi significativo em nenhuma das doze medidas. Entretanto, o valor *F* obtido para o efeito principal de sexo foi significativo nas seguintes medidas: Fluência figurativa (Círculos) $F = 9,11$, com 1 e 135 *gl*, $p < 0,001$; flexibilidade figurativa (Círculos) $F = 6,35$, com 1 e 135 *gl*, $p < 0,01$; fluência verbal (Usos Inusuais) $F = 11,99$, com 1 e 135 *gl*, $p < 0,001$; fluência figurativa (Complementação de Figuras) $F = 4,76$, com 1 e 135 *gl*, $p < 0,05$; fluência verbal (Aperfeiçoamento de Produto) $F = 9,09$, com 1 e 135 *gl*, $p < 0,001$; originalidade verbal (Aperfeiçoamento de Produto) $F = 4,45$, com 1 e 135 *gl*, $p < 0,05$.

O valor *F* obtido para o efeito de interação—Condição Experimental x Sexo—não foi significativo em nenhuma das doze medidas de criatividade.

Os alunos dos professores que constituíram o grupo experimental e de controle responderam ainda a um inventário de interesses e características que supostamente fundamentam o comportamento criativo. Os resultados obtidos neste instrumento são apresentados na Tabela 3. De modo geral, alta porcentagem dos alunos de ambos os grupos responderam positivamente aos itens denotadores de características e interesses relacionados à criatividade, como, por exemplo "eu tenho sempre perguntas que gostaria de fazer" (81,4% dos alunos do grupo experimental e 86,4% dos alunos do grupo de controle responderam afirmativamente a este item); "eu gosto de tentar novas idéias ou novas maneiras de fazer as coisas" (respondido afirmativamente por 89,8% dos alunos do grupo experimental e por 91,4% dos alunos do grupo de controle); "eu tenho muita imaginação" (respondido afirmativamente por 91,5% do alunos do grupo experimental e 87,7% dos alunos do grupo de controle).

Utilizou-se o teste qui-quadrado para análise das diferenças entre grupos neste instrumento. Diferenças significativas entre os grupos foram observadas em apenas um item, a saber: "Se eu pudesse, eu desmontaria os brinquedos para ver como eles funcionam"—67,8% dos alunos dos professores que constituíram o grupo experimental responderam afirmativamente a este item, enquanto 44,4% dos alunos dos professores do grupo de controle responderam afirmativamente ao mesmo ($\chi^2 = 6,59$, $p < 0.01$).

Tabela 3

Respostas dos Alunos dos Grupos Experimental e Controle ao Inventário de Interesses e Características

Item	Exper.		Contr.	
	N	%	*N*	%
Eu tenho sempre perguntas que gostaria de fazer.	48	81,4	70	86,4
Eu gosto de tentar e de fazer coisas difíceis.	45	77,6	55	67,9
Eu prefiro um quebra-cabeça mais fácil a um mais difícil.*	26	44,1	37	46,8
Eu sou muito curioso(a).	24	41,4	47	58,0
Eu prefiro colegas que não fazem muitas perguntas.*	38	65,5	49	60,5
Eu gosto de inventar jogos, brincadeiras, estórias ou poemas.	54	91,5	73	90,1
Quando eu quero realmente uma coisa, eu faço tudo para conseguí-la.	54	91,5	75	93,8
Se eu pudesse, eu desmontaria os brinquedos para ver como eles funcionam.	40	67,8	36	44,4
Eu gosto de tentar novas idéias ou novas maneiras de fazer as coisas.	53	89,8	74	91,4
Eu tenho muitos interesses diferentes.	41	69,5	67	82,7
Eu gosto de sonhar acordado e fazer planos para o futuro.	50	84,7	67	82,7
Eu gostaria de saber mais sobre disco voador, fantasmas e feiticeiras.	34	57,6	47	58,0
Eu nem sempre penso da mesma forma que os meus colegas.	30	51,7	38	46,9
Eu às vezes fico envolvido(a) numa determinada tarefa por muitas horas seguidas.	40	67,8	46	56,8
Eu tenho muita imaginação.	54	91,5	71	87,7
Não vejo problema em algumas vezes mudar as regras de um jogo ou de uma brincadeira.	39	66,1	52	64,2
Na minha opinião, quando se desenha o sol, ele deve ser sempre pintado de amarelo.*	47	79,7	69	85,2
Eu tenho muitas idéias boas.	56	96,6	75	93,8

*Relação inversa com maior criatividade.

DISCUSSÃO

Os resultados obtidos no presente estudo quanto à influência do programa de treinamento de criatividade nas habilidades de pensamento criativo dos alunos sugerem que este treinamento não teve um efeito significativo nas medidas de pensamento criativo utilizadas junto aos alunos dos professores que participaram do mesmo. As diferenças observadas a favor do grupo experimental em dez dentre as doze medidas de criatividade não foram estatisticamente significativas. Por outro lado, todos os professores que participaram do treinamento o avaliaram de forma positiva ressaltando, entre outros aspectos, que muito aprenderam sobre criatividade, salientando o desejo de aplicar mais atividades ligadas à criatividade e evitar comportamento e atitudes que sejam barreiras ao seu desenvolvimento e manifestação.

Várias hipóteses podem ser levantadas para explicar os resultados obtidos. Uma delas diz respeito ao momento em que foi aplicado o pós-teste. Este teve de ser aplicado imediatamente após o último encontro de criatividade, em função do término do semestre. Não foi possível aguardar algumas semanas para tal aplicação, como proposto no projeto original, uma vez que o ano letivo estava terminando naquela semana. Poder-se-ia, pois, hipotetizar que o tempo permitido aos professores para aplicar o que vinha sendo discutido nos encontros de criatividade foi muito pouco.

Uma segunda hipótese diz respeito a curta duração do treinamento. Os professores que participaram do mesmo não tinham informações sobre criatividade nem como desenvolvê-la em sala de aula. Sabe-se que muitas são as idéias errôneas a respeito de criatividade que predominam nos meios educacionais brasileiros. Uma dessas idéias é a de que apenas poucos indivíduos são criativos, visualizando a criatividade como um "dom" presente em apenas alguns poucos indivíduos privilegiados. Também comum é a idéia de que criatividade depende apenas de fatores do próprio indivíduo, subestimando-se a enorme contribuição da sociedade como um todo no processo criativo.

Grande parte dos encontros de criatividade foi direcionada para remover idéias errôneas e conscientizar os professores de que todo ser humano apresenta um certo grau de habilidades criativas, que estas habilidades podem ser inibidas e bloqueadas por um ambiente desfavorável, da mesma forma que podem ser aprimoradas através de condições ambientais favoráveis e do domínio de técnicas adequadas. Procurou-se também durante os encontros, conscientizar os professores de seu poder e influência em sala de aula e sua contribuição para o desenvolvimento de um autoconceito positivo nos alunos.

Mais tempo era necessário para o desenvolvimento de atividades práticas que poderiam ser aproveitadas em sala de aula. A insuficiência do tempo de treinamento foi inclusive destacada pelos professores ao avaliar o programa, ressaltando alguns deles que gostariam que o mesmo fosse continuado.

Pode-se destacar ainda o momento em que foi aplicado o pós-teste (no período de avaliação final do ano letivo). Sabe-se que neste momento, é comum uma ansiedade

maior entre os alunos, o que pode ter interferido no desempenho dos alunos nos testes de criatividade. Entretanto, não tínhamos escolha e não era possível adiar a aplicação dos instrumentos em função das férias letivas. Poder-se-iam destacar ainda as condições precárias e problemas enfrentados durante a aplicação dos testes de criatividade, como ausência de um local adequado para a aplicação do instrumento, interrupções durante a aplicação por professores e/ou alunos, dispensa dos alunos mais cedo, desmotivando-os a responder ao instrumento. Alguns desses problemas podem ter ocorrido com maior freqüência e prejudicado especialmente os alunos do grupo experimental, contribuindo para os resultados observados.

No inventário de interesses e características observou-se que um número significativamente maior de alunos do grupo experimental responderam afirmativamente ao item de que "se pudessem, desmontariam os brinquedos para ver como eles funcionam", denotando, pois, uma maior curiosidade.

As poucas diferenças observadas entre os dois grupos sugerem uma necessidade de um espaço maior entre o término do treinamento e a sua avaliação na amostra de alunos. Como salientado anteriormente, os professores de grupo experimental tiveram pouco tempo para aplicar o que foi discutido nos encontros de criatividade.

Destaca-se também o grande desafio enfrentado pelos professores em criar um espaço maior para o desenvolvimento da criatividade do aluno. Sabe-se que o programa adotado na escola deve ser cumprido. Este exige basicamente a memorização de fatos e conceitos, com poucas atividades para o desenvolvimento do pensamento divergente do aluno. Traços de personalidade, como iniciativa, curiosidade e independência não são tradicionalmente estimulados na escola. Noções muitas vezes incorretas são assimiladas pelo aluno. Neste sentido, poder-se-ia destacar o elevado número de alunos (79,7% dos alunos dos professores do grupo experimental e 85,2% dos alunos dos professores do grupo de controle) que responderam afirmativamente ao item "Na minha opinião, quando se desenha o sol, ele deve ser sempre pintado de amarelo". O aluno é instruído no sentido de aprender que existe apenas uma resposta correta para qualquer questão ou problema atrofiando a sua habilidade de levantar hipóteses e o seu pensamento divergente. Nota-se que, ao mesmo tempo, eles se percebem como tendo muita imaginação (que talvez seja pouco aproveitada e estimulada na escola), como tendo sempre perguntas que gostariam de fazer (e que possivelmente não estão formuladas e quando formuladas não estão sendo respondidas adequadamente), percebendo-se como portadores de muitas idéias boas (que possivelmente não estão sendo aproveitadas na escola).

Os resultados do presente estudo indicaram ainda diferenças significativas a favor do sexo feminino em várias medidas de criatividade, o que já havia sido observado anteriormente por Alencar (1975a), por Rodrigues e Alencar (1983) e por outros pesquisadores como Burgess (1971) e McDaniel (1974).

Uma possível explicação para tais diferenças diz respeito à menor motivação pelo que ocorre na escola por parte dos alunos do sexo masculino e especialmente por aqueles de status sócio-econômico baixo. Estes aspiram iniciar-se mais cedo no mercado de trabalho e são também mais pressionados pela família neste sentido. Tanto

o programa escolar como as atividades desenvolvidas em sala de aula estão bem distantes dos interesses desta amostra de alunos que tendem a demonstrar maior apatia e desatenção, apresentando, na opinião dos professores, um maior número de características consideradas indesejáveis. Ao serem escolhidos para responder aos testes de criatividade, é possível que estes sujeitos tenham reagido da mesma forma como o faziam diante de qualquer tarefa proposta por seus professores—menos motivados e empenhando-se menos na realização das tarefas propostas.

Para entender melhor tais diferenças, seria relevante que estudos fossem realizados com amostras de alunos de ambos os sexos, manipulando as condições de aplicação do instrumento, como, por exemplo, aplicando-o individualmente ou oferecendo incentivos àqueles que apresentassem um melhor desempenho.

REFERÊNCIAS

Alencar, E.M.L.S. (1974). *A study of creativity training in elementary grades in Brazilian schools*. Tese de doutorado, Purdue University, USA (a).

Alencar, E.M.L.S. (1974). Avaliação da criatividade do aluno por professores. *Revista Interamericana de Psicología/Interamerican Journal of Psychology, 8*, 219–224 (b).

Alencar, E.M.L.S. (1974). Um estudo de criatividade. *Arquivos Brasileiros de Psicologia Aplicada, 26*, 59–68 (c).

Alencar, E.M.L.S. (1975). Efeitos de um programa de criatividade em alunos de 4ª e 5ª séries. *Arquivos Brasileiros de Psicologia Aplicada, 27*, 3–15.

Alencar, E.M.L.S. (1975). Relação entre o nível de criatividade do professor e de seus alunos. *Revista Brasileira de Estudos Pedagógicos, 61*, 376–380.

Alencar, E.M.L.S. (1984). Características psicossociais de alunos mais e menos criativos. *Revista Interamericana de Psicología/Interamerican Journal of Psychology, 18*, 87–100.

Alencar, E.M.L.S. (1985, January/February). Creativity in Brazilian schools. *G/C/T*, pp. 15–17.

Alencar, E.M.L.S. (1986). *Psicologia da criatividade*. Porto Alegre: Artes Médicas Sul.

Burgess, W.V. (1971). The analysis of teacher creativity, pupil age and pupil sex as sources of variation among elementary pupils' performance in pre and post-tests of creative thinking. *Dissertation Abstracts International, 32*, 747–748 (University Microfilms Nº 71-20, 779).

Covington, M.V., Crutchfield, R.S., & Davis, L. (1966). *The Productive Thinking Program*. Berkeley: Education Innovation.

Davis, G.A. (1975). In frumious pursuit of the creative person. *Journal of Creative Behavior, 9*, 75–86.

Feldhusen, J.F., Treffinger, D.J., & Bahlke, S.J. (1970). Developing creative thinking: The Purdue Creativity Program. *Journal of Creative Behavior, 4*, 85–90.

Gordon, E.J.J. (1971). Synectics. Em G.A. Davis & J.A. Scott (Eds.), *Training creative thinking*. New York: Holt, Rinehart & Winston.

Guhur, V.M. (1976). *Efeitos da modelação e instruções sobre o comportamento criativo em adolescentes*. Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Houston, J. (1977). The psychenaut program: An exploration into some human potentialities. Em S.J. Parnes, R.B. Noller & A.M. Biondi (Eds.), *Guide to creative action*. New York: Charles Scribner's Sons.

Maia, F.R.R.R. (1975). *Efeito do reforço no desenvolvimento de criatividade verbal em escolares*. Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Marin, A.J. (1976). *Educação, arte e criatividade*. São Paulo: Pioneiras.

McDaniel, E.R. (1974). An investigation of the relationship between creativity and socio-economic status, race and sex of seventh-grade pupils. *Dissertation Abstracts International, 34*, 3997–3998 (University Microfilms Nº 73-32, 099).

Osborn, L.E. (1963). *Applied imagination*. New York: Charles Scribner's Sons.

Petrosko, J.M. (1978). Measuring creativity in elementary school: The current state of the art. *Journal of Creative Behavior, 12*, 109–119.

Rodrigues, C.J.S., & Alencar, E.M.L.S. (1983). Um estudo de pensamento criativo em alunos do primeiro grau. *Arquivos Brasileiros de Psicologia, 35*, 113–122.

Torrance, E.P. (1966). *Torrance Tests of Creative Thinking. Norms—Technical Manual* (Research Edition). Princeton: Personnel Press.

Torrance, E.P. (1972). Predictive validity of the Torrance tests of creative thinking. *Journal of Creative Behavior, 6(4)*, 236–252.

Treffinger, D.J. (1979). Fostering independence and creativity. *Journal of the Education of the Gifted, 3*, 215–224.

Wechsler, S. (No prelo). Efeitos do treinamento de criatividade em crianças bem-dotadas e regulares do primeiro grau. *Revista Brasileira de Estudos Pedagógicos*.

Recebido em 17 de dezembro de 1986
Revisão recebida em 5 de outubro de 1987

AUTORES

EUNICE MARIA LIMA SORIANO DE ALENCAR. *Instituição:* Universidade de Brasília, Brasília, Distrito Federal, Brasil. *Título:* Professor Adjunto. *Diplomas:* Licenciatura em Psicologia, 1966, e Formação de Psicólogo, 1967, Universidade Federal de Minas Gerais, Brasil; Master of Science, 1970, e Ph.D., 1974, Purdue University, U.S.A. *Especialização:* Psicologia do desenvolvimento e escolar. *Endereço:* SHIS, QL 10, Conjunto 6, Casa 14, 71600, Brasília, DF, Brasil.

DENISE DE SOUZA FLEITH. *Instituição:* Universidade de Brasília, Brasília, Distrito Federal, Brasil. *Título:* Aluna de Mestrado. *Diplomas:* Licenciatura em Psicologia e Formação de Psicólogo, 1985, Universidade de Brasília. *Especialização:* Psicologia do desenvolvimento e escolar. *Endereço:* Universidade de Brasília, Departamento de Psicologia, Campus Universitário, Asa Norte, CEP 70910, Brasília, DF, Brasil.

LOURDES AKICO YONAMINE SHIMABUKURO. *Instituição:* Universidade de Brasília, Brasília, Distrito Federal, Brasil. *Título:* Bolsista de Iniciação Científica (Conselho Nacional de Desenvolvimento Científico e Tecnológico). *Diploma:* Licenciatura em Psicologia, 1976, Faculdade de Filosofia, Ciências e Letras de Santos. *Endereço:* SQS 204, Bloco D, Apartamento 308, CEP 70234, Brasília, DF, Brasil.

MÁRCIA DE ANDRADE NOBRE. *Instituição:* Universidade de Brasília, Brasília, Distrito Federal, Brasil. *Título:* Bolsista de Iniciação Científica (Conselho Nacional de Desenvolvimento Científico e Tecnológico). *Diploma:* Graduação em Psicologia, 1987, Universidade de Brasília. *Endereço:* SHIN, QI 1, Conjunto 7, Casa 14, CEP 71500, Brasília, DF, Brasil.

EXTENDED SUMMARY

Effects of a Creativity Training Program for Elementary School Teachers on their Pupils' Creative Thinking Abilities

EUNICE MARIA LIMA SORIANO DE ALENCAR
DENISE DE SOUZA FLEITH
LOURDES A. SHIMABUKURO

and

MÁRCIA DE ANDRADE NOBRE

The purpose of the study was to evaluate the effect of a creativity training program for elementary school teachers on their students' creative abilities and characteristics related to creativity, such as curiosity, independence, and perseverance. The sample were 42 fourth-grade public school teachers of a satellite city in the Federal District, Brazil. Teachers were assigned to two different conditions: creativity program and control. Prior to the beginning of the creativity program two verbal (Unusual Uses and Product Improvement) and two figural tests (Circles and Picture Completion) of the Torrance Tests of Creative Thinking were administered to a random sample of 140 students of these elementary school teachers. After the three-month creativity training program, these tests were again administered to the same sample of students. A questionnaire was also administered to the students to assess personality characteristics related to creativity, such as flexibility, curiosity, and persistence. The creativity program was designed to instruct teachers in techniques and exercises that might be used in the classroom to foster creative abilities in the students. Once a week for ten weeks, teachers from the experimental group met with the instructor to be trained in the production and evaluation of ideas and in techniques that might be used in school to develop children's creative thinking abilities and to be given information about creative thinking, personality traits relative to creativity, and the several obstacles to the development of creativity in school and society.

Analysis of covariance and chi-square were used to analyze the data. No significant differences were observed between the two samples of students on the Torrance Tests of Creative Thinking, although higher scores were observed for figural fluency; flexibility, and originality for the Circles and Picture Completion tests; verbal fluency, flexibility, and originality for Product Improvement; and verbal flexibility for Unusual Uses, among students of the teachers who participated in the creativity training program. Sex differences were found on six variables, with female students' performance exceeding male students on figural fluency and flexibility on the Circles test, figural fluency on the Picture Completion test, verbal fluency on the Unusual Uses and Product Improvement tests, and verbal originality on the Product Improvement test, regardless of their teachers' participation in the creativity program. It was also

observed that most students answered positively to items that indicated curiosity, such as "I always have questions that I would like to ask" or "If I have permission, I would take toys apart to see how they work;" persistence, such as "When I want something, I try my best to get it," and imagination, "I have a lot of imagination." A significant difference in favor of the students whose teachers participated in the creativity program was observed in one item of the questionnaire that measured curiosity. Inadequate conditions during the administration of the Torrance tests to the students (no school has a vacant room for this purpose) and the time of the evaluation (immediately after the completion of the program) may explain the small differences observed among the students of the two groups of teachers.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp.72-89

Sources of Chronic Rural Poverty: Behavioral Perspectives

SHARON ALISA LOBEL

Harvard University
U.S.A.

Diverse behavioral approaches have been proposed to explain endemic rural poverty in developing countries. The *traditional* peasant perspective pinpoints aspects of personality (e.g., low need for achievement) that predispose peasants to resist adoption of technological innovations. Rural development strategies implied by this approach include targeting persons predisposed to modernity and educating peasants to adopt modern attitudes. The view of peasants as *risk averse* focuses on the farmers' rational assessments of the risk involved in new technology adoption. The implied strategy for rural development is reduction of risk through technological innovations (e.g., drought resistant seeds). The *helpless* peasant perspective suggests that peasants who perceive lack of personal control over their agricultural production are less likely to engage in behaviors designed to improve their situation. The implied rural development strategy involves training peasants to identify internal and external control factors affecting production in predictable ways. Each of these behavioral perspectives is described and critiqued. Whereas these approaches largely ignore the environment of the poor farmer, and therefore are of limited utility, purely structural explanations are not sufficient either. Improved management techniques and program design based on synthesis of local world views with those embodied by the intervention represent two areas in which behavioral strategies may be quite effectively employed.

The recent earthquake in Mexico City caused international attention to rivet upon the world's largest city. A crisis was certainly at hand. Yet, in far less dramatic circumstances, a crisis has been at hand for quite some time. According to a Mexican

The author gratefully acknowledges the comments of Aroldo Rodrigues and two anonymous reviewers.

Copyright © 1987 by Sharon Alisa Lobel. All rights reserved.

government study completed in 1979 (Oficina de Asesores del Presidente, 1980), nearly 70% of rural workers were unemployed or underemployed, in comparison with 40% in 1970. Almost 90% of the rural population suffers from calorie and protein deficiency, at a level which for the majority of rural areas has either declined or remained stable since 1959. These statistics no doubt account for the flight of 1,000 rural poor to Mexico City each day. The rural poor are running from poverty and agricultural stagnation, from problems that are hardly unique to Mexico.

The backbone of agricultural production in developing countries continues to be the village peasant. But why are peasants not benefiting from 20th century agricultural science? Why are rural development efforts not achieving their targets? And what is to be done? This article reviews the state of the art with respect to behavioral answers to these questions. The three principal behavioral explanations—the traditional peasant, the risk averse peasant, and the helpless peasant—imply different strategies for rural development, and these strategies are detailed, along with a critique.

THE TRADITIONAL PEASANT

> In many parts of the world we find cultures adhering to the belief that man has no causal effect upon his own future or the future of the land; God, not man, can improve man's lot. . . . It is difficult to persuade people [with these beliefs] to use fertilizers, or to save the best seed for planting, since man is responsible only for the performance, and the divine for the success, of the act. (Mead, 1955, p. 185)

This excerpt from Margaret Mead's manual for change agents, *Cultural Patterns and Technical Change*, is characteristic of the conception of a tradition-bound peasantry, an idea most popular among social scientists during the 1950s and 60s. Basically, the position holds the poor to be poor because they do not want to give up their old ways, in some cases actively resisting change.

The poor belong to traditional societies, which represent the "backward" end of a traditional/transitional/modern continuum of societies (Lerner, 1958). Change, or escape from poverty, will come about slowly as a result of culture contact. Culture contact serves as a stimulus for those people within traditional societies, who are *predisposed* to change, to take advantage of the new opportunities presented to them. People so predisposed have been variously referred to as mobile personalities (Lerner, 1958), modern personalities (Inkeles & Smith, 1974), high-need achievers (McClelland & Winter, 1969), or innovators (Rogers, 1962).

An example of this approach is Inkeles and Smith's study of six countries, carried out in the 1960s and published as *Becoming Modern* (1974). The authors developed an Overall Modernity scale, operationalizing characteristics of the modern personality, such as openness to new experience, awareness of diverse attitudes and opinions, belief

in individual choice and efficacy, and respect for subordinates. Endorsing items such as "Good planning and hard work are more important for the future of (Argentina) than luck and God's help" would qualify as a modern answer. Six thousand men — cultivators, urban non-factory workers, and urban factory workers —were interviewed in Argentina, Chile, East Pakistan (now Bangladesh), India, Nigeria, and Israel. The authors chose the factory setting because the modern factory "exemplifies efficacy. . . . It affirms man's capacity through organization and the harnessing of mechanical power to transform nature to suit his needs" (p. 158). The results indicated that the longer one worked in a factory, the higher one's score on the Overall Modernity scale. The authors concluded that without the spirit of modernity, "neither foreign aid nor domestic revolution can hope successfully to bring an underdeveloped nation into the ranks of those capable of self-sustained growth" (p. 315).

In a similar vein, McClelland (1977) observed:

> Certain human characteristics — especially . . . need for Achievement — seem to precede and to promote more rapid economic development and modernization. . . . Educating men to have this characteristic led them to become more active businessmen who created the kinds of firms which eventually could serve as schools of modernization. (p. 45)

The characteristics of belief in efficacy and need for achievement are also implied in Rogers' (1962) list of attitudes held by agricultural innovators. Examples of personality characteristics he describes are: (a) dissatisfaction with present levels of production or consumption, (b) confidence that increased production is possible, (c) willingness to experiment even in the face of possible failure, (d) confidence in agents of change, (e) sense of importance of time and planning, and (f) readiness to make independent decisions.

The traditional vs. modern personality orientation has its macro corollary in the dual economy model of underdevelopment (Jorgenson, 1969). This theory assumes the coexistence of two more or less autonomous sectors within a given national economy— one modern, the other backward. The modern sector typically is engaged in export production from large agricultural units and agricultural industry, while the backward sector is characterized by subsistence farming, low-level technology, and archaic social organizations. A malevolent environment is not implied. Rather in Dillon's (1979) poker analogy, "It is not that the cards have been stacked against the small farmer, but just that the deal has not turned out too favorably. If the game goes on long enough, things will get better" (p. 71).

Migdal (1974) summarizes the essential ingredients of the traditional peasant perspective, with respect to social and technological change, as follows: (a) The benefits of the modern are far superior to the traditional (society, technology, personality, etc.); (b) individuals are free from external constraints that might prevent their own choice of change behaviors; and (c) those individuals who select the new are rational and are optimizers, and those individuals who do not accept the modern fail to do so because of wrong or nonrational values.

Implications for Rural Development Strategy

If we believe the source of chronic poverty to be the absence of particular personality traits, characteristic of persons in modern, industrialized societies, then the implications for strategy can take two forms. First, one can assume a laissez-faire attitude. In other words, those who are ready to benefit from new technology will do so, and those who are not ready, will eventually benefit as a result of trickle-down. Modernization of attitudes will occur at its own unhurried pace, as a result of culture contact, slow diffusion of improved techniques adopted from the modern culture by innovators predisposed to modernity, eventually benefiting the laggards (Rogers, 1962) in the backward economy.

The second strategy that might be implied by this approach is slightly more active and is referred to by Cancian (1977) as the *education tradition*. If poor people stay poor because they have the wrong attitudes, then we must educate them to be more like us in their understanding of the world. Recent advocates of the education tradition include Galbraith (1979) and McClelland.

Critique

The major limitation of this form of analysis is that it simply discounts the environment of the poor farmer, which often severely restricts the choice of change behaviors. Let us examine the case in which a bank uses land title as collateral for the provision of credit. Should there be a drought, to take the simplest example, the farmer stands to lose a great deal. The large landowners, in such a situation, are likely to survive, not because of better management of resources, or more openness to change, but principally because they have a reserve to fall back on. In these cases, personalities and limited knowledge are not the key obstacles to change. Environmental factors may provide more compelling explanations. The behavioral analysis of the traditional peasant might be restated as lack of knowledge about some forms of technology, rather than predispositions to resist all forms of technology-based change. With a focus on managing the implementation of specific technologies, the psychologist cannot avoid examining the environmental factors that favor or constrain technology adoption.

THE RISK AVERSE PEASANT

The view of peasants as economically rational and risk averse is currently quite popular among rural development planners and politicians. This perspective represents a mini-paradigm shift (or a mini-discipline shift) away from viewing farmers as poor because they are tradition-bound, fatalistic, and irrational, towards viewing farmers as local experts, open to change at acceptable levels of risk, and rational in their choices — which nonetheless leaves them poor.

The broader theoretical orientation had its roots in Schultz's (1964) *poor but efficient* model of peasant behavior. According to this model, small farmers operate in a relatively static economic, technological, and cultural environment to which they have adapted, and in which they operate efficiently in economic terms. Change will come about through a system of incentives (e.g., profitable new technology or education). It is not a matter of slow change in attitudes as analyzed in the portrait of the traditional peasant. Rather, with the proper incentives, that is to say, if a technology is not too risky, this model assumes that change (technology adoption) will occur rapidly. As with the literature on the traditional peasant, a malevolent environment is not implied.

Although individual farmer decisions may be more or less risk averse, models of farmer behavior generally group farmers together. The questions that are posed take the form, "Do small farmers make choices so as to maximize profits or to avoid risk?" "Given this crop, these input costs, and this probability distribution of benefits and losses, at what level of profitability or risk will farmers choose to adopt?" Investigators model the farmer's decision-making and risk-taking behavior under the different conditions of interest.

Hypotheses derived from describing the farmer's behavior as rational and risk averse include the subsistence objective (Sanders & de Hollanda, 1979), which states that small farmers will first satisfy subsistence requirements and then maximize income, constrained by the highest risk level they are willing to take. In theory, in order to assess the perceived riskiness of the technology and predict whether farmers will be willing to adopt it, historical data about yield and price variation can be modeled, with farmers subjectively weighting the probability distribution.

A variant of the subsistence objective is the safety-first hypothesis (Day, Aigner, & Smith, 1971), which states that the primary objective of small farmers is to achieve a minimum income goal. Thus, nonadoption of more profitable crops is explained as rational behavior, given the farmer's safety-first objective. For example, crop diversification (i.e., subsistence – cash crop mixes) guarantees the farmer a minimum income under any weather–price conditions.

The investigation by Zandstra, Swanberg, Zulberti, and Nestel (1979) of the Caqueza project in Colombia serves as one example of risk assessment in the field. The project faced the problem of nonadoption of a new corn technology. The investigators attempted to identify whether the constraining factor was the riskiness of recommendations. They classified risks as production-, marketing-, and institution-related, and they calculated that the recommendations were too risky. A program of shared risk was instituted, which guaranteed that if production was below a certain level per hectare, farmers would not have to pay for the material inputs received for the new technology. Beyond this level, the quantity produced would be divided among the farmer and the plan organizers (a local producer cooperative). As might be expected, this plan required extensive field inspections and accurate yield estimates beyond the technical capacity of the staff. As a result, while the farmer's earnings increased by $43/hectare, the cooperative suffered a loss, amounting to the corn withheld by farmers who failed to fulfill their obligations.

Implications for Rural Development Strategy

If we accept the view that farmers stay poor because they rationally make decisions to minimize risk, then it follows that farmers will become less poor in response to strategies which reduce the amount of risk associated with adoption, for example, crop insurance and the development of disease and drought resistant seeds. The focus is upon technological solutions, and risk must be largely understood in terms of quantifiable entities—input costs, yield, market price—which can be currently measured or averaged over time.

Critique

Critiques of this approach can be divided into those that question how risk should be defined and those that question how risk should be measured.

Definition. Risk and rationality have become catch-all terms that cover a range of constraints facing the farmer—soil conditions; disease susceptibility; competition for cash, labor, and water; timely delivery of inputs; availability of transport and storage; skills, equipment, and technical assistance requirements; attitudes and dietary preferences—not to mention corruption and political control. Furthermore, as one economist openly admits, "Since it is often difficult to criticize official recommendations and since it is no longer fashionable to assume farmers are either irrational or lazy, the hypothesis of risk aversion provides a convenient way of resolving the apparent paradox that farmers are rational but inefficient" (Dillon, 1979, p. 59).

What economists mean by risk has not met with agreement (Roumasset, 1977; Sanders & de Hollanda, 1979), though for the most part, they do not mean risks associated with institutional and social constraints. One recent study (Binswanger, 1980) showed, however, that it was precisely differences in external constraints such as access to credit, marketing, and extension that explained differences in farmer adoption within regions of similar climate and soil conditions. Across farmers in the same region, the spread in willingness to take risks was found to be small. Binswanger and Sillers (1983) assert that, in general, small farmers are *not* less willing to take risks than large farmers; rather, the same external conditions translate into higher risks.

Measurement. In the field, it may not always be possible to obtain estimates for adequate modeling of farmer decision-making, due to lack of historical data or the difficulties in obtaining reliable probability estimates from farmers. More important, estimates of risk need to take into account the risk faced by individual farmers (not average farmers) and the risk faced by groups of farmers in a social setting. The need to examine individual farmers stems from the greater sensitivity of new farming techniques to inevitable variation in soil and climate conditions. Results obtained on irrigated test plots will differ considerably from those obtained in the "real world" on rainfed land by a factor of about eight to one (Cancian, 1977). It is clear that for farmers, to whom risk and safety-first mean the difference between survival and starvation,

nothing short of individual estimates will be practical. This is not to say that average estimates cannot suffice for farmers who are able to absorb risk.

Further understanding risk to include variation and uncertainty in the farmer's social environment is also desirable if the risk construct is to be meaningful. Credit and crop insurance are only valid buffers if the bureaucracy that provides them is more reliable than not and if incentives to cheat can be controlled. This means that risk reduction and assessment must take into account administrative mechanisms and performance criteria as well as the more common production variables.

If we return to the paradox of why farmers stay poor even though they make rational and efficient decisions, it is evident that it is not enough to say that farmers are poor because they are risk averse, unless the principal factors contributing to risk level can be specified with precision.

THE HELPLESS PEASANT

The psychologist Leonard Doob wrote that basic to any conception of the future as a realm of potential change "must be the belief that men themselves—not their ancestors, not fate, not nature, not other men — are able to control their own destinies . . . for men everywhere are not likely to seek change unless they believe that change is possible" (1960, p. 227). The psychological belief in control over one's destiny has been identified as one of several features in measures of modernization, as discussed earlier. Here, we are examining the control belief, not as characterizing a behavior trait, but as an expectancy that is created and subject to change through the learning process.

According to social learning theory (Rotter, 1954), people differ in the extent to which they expect to have control over their own lives. Those who believe that they have control over most of the events in their lives are said to have an *internal locus of control*. Those who believe that their lives are controlled by factors other than their own actions are said to have an *external locus of control*. Social learning theory further suggests a relationship between expectancies of external control and reduced levels of motivation in the pursuit of valued goals:

> Whereas fatalism or external control beliefs are associated with apathy and withdrawal, the holding of internal control expectancies presages a connection between an individual's desires and his subsequent actions. As such, locus of control can be viewed as a mediator of involved commitment in life pursuits. If one feels helpless to affect important events, then resignation or at least benign indifference should become evident with fewer signs of concern, involvement and vitality. (Lefcourt, 1976, p. 152)

A social learning theorist would propose that the rural poor stay poor because they believe that external forces — ancestors, fate, nature, other people — control whether

they obtain increased production, income, or access to opportunities associated with new technology.

A person develops control expectancies on the basis of past experiences. Concrete experiences of personal control over outcomes generalize to expectations of internal control over similar events in the future. Likewise, the perception that external factors are responsible for specific events will generalize to expectations of external control over similar behavior–outcome sequences in the future. By implication, individual differences in outcome or reinforcement histories will explain some of the differences in how individuals learn from experience and the corresponding variations in behavioral choices they make.

A number of studies have demonstrated a relationship between the cognitive processes that are essential to learning and the expectancies of internal or external control. Specifically, as reviewed in Lefcourt's (1976) chapter on locus of control and cognitive activity, an internal locus of control expectancy has been linked with greater attention, deliberation, inquisitiveness, and utilization of information relevant to goal attainment. Most of the research reviewed uses laboratory manipulations and Rotter's 23-item forced-choice scale to assess locus of control.

Other studies have examined the relationship between locus of control and behaviors aimed towards improving one's situation. Most often cited are Strickland's (1965) correlational study of active engagement in civil rights groups among U.S. Black college students, and Gore and Rotter's (1963) use of locus of control to predict U.S. Black college students' verbal commitment to participate in protest movements for civil rights. In both cases, internals were significantly more active or more likely to indicate willingness to participate in civil rights activities than externals.

The locus of control concept has also been studied in the context of rural development. The relationship between control beliefs held by Punjab farmers and productivity was investigated by Singh (1983) in a region with very favorable conditions (fertile soil, irrigation). Singh divided the 320 farmers in the sample into two groups based on per acre farm output between 1978 and 1980. A number of factors significantly distinguished the two groups. Incentives to productivity included upward striving, punctuality, and tolerance for work pressure. External control beliefs (chance, fate, or luck), as measured with a locus of control scale, and an "interest in mating" were cited as factors associated with retarding farm productivity.

Implications for Rural Development Strategy (I)

The general thrust of this research, as described so far, would suggest a strategy similar to that under the traditional peasant section—namely, training for internality and efficacy. The effectiveness of internality counseling in the rural setting was tested with a sample of small coffee producers in Venezuela (Díaz, Rengifo, & Romero García, 1985). A series of directed informal and formal meetings centered on activities that would elicit the farmers' characteristic motivational patterns and control beliefs.

Comparison with a control group of farmers revealed an increase in internal control beliefs as a consequence of the intervention. The trained farmers subsequently achieved the satisfaction of two demands that had been placed upon the government, namely, assignment of a new coffee technician and nurse. The trained farmers also demonstrated initiative in constructing a storage bin (O. Romero García, personal communication, August 20, 1986).

Critique

Critiques of this analysis and its implications for strategy focus upon (a) failure to account for veridical external control expectancies, and (b) the generalizability of training for internality.

Veridicality. Expectancies of control are subjective perceptions that may more or less veridically reflect actual contingencies between behavior and outcomes. In the case of poor people and groups suffering from historical discrimination, it is likely that expectancies of external control are more veridical than not. For example, Putterman (1985) argues that voluntary agricultural collectivization in Tanzania and China encountered problems not because collectives in themselves are ineffective, but because obstacles to self-determination prevent peasants from being free agents. In the cases cited, these obstacles included externally made decisions about collective leadership, size and location of collectives, and paternalistic, authoritarian attitudes towards peasants.

Some studies have shown that perception of external control contingencies may be a function of the perceiver's intelligence, education, or socioeconomic condition. For example, Battle and Rotter (1963) found that lower-class U.S. Black children of high IQ were more likely to believe in external control than lower-class U.S. Black children of low IQ. A study of perceived causes of poverty among Europeans (Commission of the European Communities, 1977) found that the better educated subjects tended to blame societal causes (e.g., injustice in society), while the less educated subjects blamed individual causes (e.g., laziness). Recently a survey of perceived causes of poverty among Barbadian and Dominican adolescents (Payne & Furnham, 1985) showed that while both groups stressed societal over individualistic or fatalistic explanations, the Barbadian adolescents showed more concern about social inequalities than the Dominicans. The authors suggested that this may be a consequence of greater public awareness among Barbadians.

Evidently control may be exercised by powerful members of the society in predictable ways over persons with well-defined attributes, such as skin color or low income. Yet, the concept of external locus of control, as originally formulated by Rotter and as measured in many studies since, does not distinguish undifferentiated chance factors from identifiable external sources of control. For this reason, Levenson (1981) has revised the scale to measure three types of expectancies: internal control, control by chance, and control by powerful others. This revision coincided with a call for

development of new locus of control scales to measure control expectancies in specific settings and to reduce the growing trend to view locus of control as a trait with similar behavioral consequences for all groups of people (Lefcourt, 1981).

Generalizability of training for internality. It has been shown that external control expectancies do not necessarily imply reduced motivation or fatalism for all classes of people. For example, Levenson's (1981) research on sex differences in expectancies of control indicated that women seem to achieve more success when they view powerful others as playing an important role in their lives, while the same is not true for men. Similarly, Gurin and Katz (1966) reported that U.S. Black job trainees who blamed the system for their previous failures were more likely to find jobs than those who blamed themselves.

Since different groups hold different control beliefs with varying levels of adaptive success, the effectiveness of counseling or training designed to increase internality should depend upon the characteristics of the group that is being trained. Indeed, Roueche and Mink (1976) reported a significant gender difference in response to internality counseling: Female college students were less internal to begin with and were less likely to increase in internality over time as a result of counseling. Furthermore, training for internality must coincide with a realistic assessment of environmental opportunities. Otherwise the subject could theoretically be led to blame himself or herself, with a corresponding decrease in self-esteem, for events that require more than a change in individual expectancies to produce favorable outcomes.

In some situations, persons may be inclined to relinquish control as a means of adaptively avoiding more aversive outcomes. Miller (1980) suggests that these conditions may involve the presence of an external stable factor perceived to be able to reduce the maximum amount of harm the person expects to experience. An example from the rural setting (Scandizzo, 1979) concerns the relationship between sharecroppers and landlords in Brazil:

> Because of the high interest rates; the comparatively low prices for the product; and the wild oscillation of production due to pests, diseases and periodic droughts, the sharecropper is often caught in a relation of perpetual indebtedness and dependence on the landlord. This relation is, however, not forced upon the farmers but is accepted by them as a sort of social form of insurance against the extreme risks of the environment. (p. 124)

This dismal description highlights the need for an analysis of the farmer's situation that accounts for *why* an individual maintains a particular pattern of control beliefs in addition to detailing what the beliefs are and how they might ideally be changed.

Implications for Rural Development Strategy (II)

Expectancy change can come about through cognitive restructuring or through change in the objective environment. In the rural setting, expectancies of control by

powerful others, rather than expectancies of control by chance factors or expectancies of internal control, may be most realistic for poor farmers. If poor farmers hold expectancies of control by chance, learning that valued reinforcements are controlled by powerful others in predictable ways may be a necessary but not sufficient condition for the ability to bargain for and achieve immediate and longer-range benefits from new technology. Whereas risk aversion and traditional peasant analyses focus upon technology and personality as the catalysts for behavior change, socially based expectancies of control point explicitly to the role of powerful persons and institutions in promoting, supporting, or inhibiting change. Nonetheless, changing control expectancies represents merely a small portion of the rural development process and is likely to be ineffective without other supporting strategies.

CONCLUSION

As the review of behavioral perspectives above has demonstrated, neglect of situational factors seriously undermines the power of behavioral approaches to explain and solve the problem of chronic rural poverty. Thus it is not surprising that these approaches are adhered to by a minority of scholars of peasant studies. A survey of the articles published in the *Journal of Peasant Studies* from October 1981 through October 1984 revealed that none of the 47 regular feature articles presented a primarily behavioral approach. This finding may, of course, stem from selection processes, since the journal clearly favors historical accounts and Marxist or neo-Marxist analyses of class relations. Interestingly, in a special section of the journal entitled "Peasants Speak," which appeared only twice during the three years that were surveyed, Goodell (1984) presented a diary of a Filipino farmer reporting his experiences with a new rice technology over an eight-year period. The diary revealed the erratic performance and high cost of the new rice seed, clearly very risky for the farmer. The farmer also highlighted control of his production by external factors. For example, the credit program was linked to purchase of seeds from a specific higher cost supplier. Goodell concluded that "Mang Lino's notebook suggests that the most dangerous aspect of the new rice technology . . . [is] the way it has been set forth to farmers as full-blown, externally dictated and complete" (p. 265). She urged that technology be developed by, or in close interaction with, those most likely to use it, and further, that rural development professionals should be accountable to beneficiaries rather than bureaucrats in capital cities.

Case studies, such as these, appeal time and again to the need for programs that are tailored to local conditions. For this reason, purely structural explanations of chronic rural poverty are *not* guaranteed to generate strategies that will function more effectively than purely behavioral approaches. Moris (1981), for example, suggests a few of the problems that may impede effectiveness of a purely socialist rural development effort: (a) the bureaucracy may be monolithic; (b) market mechanisms

may be abolished before administrative capability for planning has been established; (c) the party does not always reflect peasants' interests; (d) promotions may be made for fervor, not expertise; and (e) principles such as "workers' solidarity" do not indicate specific implementation procedures, namely, how the rural economy will function after the revolution.

According to Moris, rural development efforts, whether based on socialist or capitalist ideologies, fail because projects are designed externally and superimposed; they are not attuned to farmers' incentives; they are not targeted towards specific clientele; and personnel are poorly managed. These "human" factors are clearly powerful determinants of development, alongside broader social, structural, political, and economic conditions.

Evidently, the complexity of rural development demands communication among professionals from different disciplines. Although the behavioral approaches presented thus far have been shown to be inadequate, there is a role for social psychologists and other social scientists who are concerned with behavioral solutions, if not explanations, for rural poverty. Two examples of directions that might be pursued are (a) better management of rural development and (b) greater identification and utilization of local world views.

Management of Rural Development

Poor farmers are clearly at the mercy of external forces. Even a rural development program that uses participation approaches can be undermined by powerful local elites or by a proliferation of social services without guaranteeing a productive fiscal base to support them. Grass-roots programs may do well until they expand and come into conflict with existing bureaucratic structures. Therefore, data need to be collected on behavioral variables that will shape administrative capabilities, such as nature and strength of local leadership, degree of cooperation among groups that will be involved in the program, degree of alienation of various sectors, reward distribution in the community, and attitudes towards important aspects of the intended program. According to Moris (1981) effective management of rural development can be built upon these data and further enhanced by appropriate organizational behavior, such as simple supervision and a short chain of command, a core staff that has already worked in the region, efforts to minimize farmer risk, building the project into the local administrative structure, and making sure that staff are adequately trained, motivated, and supported.

Utilization of Local World Views

A key element of successful management practice is the design and adaptation of the program to local conditions. The consequence of taking this oft-cited dictum

seriously is that we move from an etic, usually Euro-American, vision of rural development to one that is emic and perhaps more able to avoid some common pitfalls. An example of such an approach is given by Ramirez (1983). Ramirez, like Moris, cites the importance of on-site behavioral observations. He goes a step further in suggesting behavioral techniques, such as role-playing, to train beneficiaries how to interact with change agents and vice versa. This is a truly interactionist approach in which individuals learn how best to interface with institutions.

Behavioral observations will need to include an assessment of the peasants' world view. If done without ethnocentric biases, alternative data collection methods (e.g., life histories) and alternative development outcomes are likely to emerge. For example, as detailed by Ramirez, the mestizo world view is one that encourages openness to diversity. Therefore, an acceptable development alternative, which is based on the local mestizo world view, would be a synthesis of cultural contributions and not a mere abandonment of traditional ways in favor of adoption of a "superior" technology.

Effective interventions based on local world views have also been described by Freire (1970). He instructed literacy teachers to discover what words designate issues of immediate import to peasants (e.g., access to a well or compound interest on debts) and to use these words as the foundation for literacy training. Individuals so taught not only learn to read, but also grow in social awareness and impulse to political action. By analogy, issues that are salient to farmers could be assessed, and new farming techniques might be taught as merely one means of improving production, an adjunct to farmer-inspired measures for eliciting appropriate support from change agents in handling inevitable conflicts of interest. Empirical research to determine the best nonformal education methods to use with specific farmers in a particular region is another task for behavioral scientists (cf. Laosa, 1976).

These two behavioral strategies — improved management and utilization of local world views — represent advances over the earlier reviewed approaches in that: (a) they do not assume that the modern is superior to the traditional; (b) risk factors will be minimized as programs are designed based on local conditions; and (c) comprehensive interventions examine but are not limited to identifying the key external factors that control production.

REFERENCES

Battle, E., & Rotter, J.B. (1963). Children's feelings of personal control as related to social class and ethnic groups. *Journal of Personality, 31*, 482 – 490.

Binswanger, H. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics, 62*, 395 – 407.

Binswanger, H., & Sillers, D. (1983). Risk aversion and credit constraints in farmers' decision-making: A reinterpretation. *Journal of Development Studies, 20*, 5 – 21.

Cancian, F. (1977). Can anthropology help agricultural development? *Culture & Agriculture, 2*, 1–7.

Commission of the European Communities (1977). *The perception of poverty in Europe*. Brussels: Author.

Day, R.H., Aigner, D.J., & Smith, K.R. (1971). Safety margins and profit maximization in the theory of the firm. *Journal of Political Ecomony, 79*, 1293–1301.

Díaz, C., Rengifo, E., & Romero García, O. (1985). *Cambios hacia la internalidad en pequeños productores agrícolas*. Unpublished manuscript, Universidad de los Andes, Facultad de Medicina, Laboratorio de Psicología, Mérida, Venezuela.

Dillon, J.L. (1979). Broad structural review of the small farmer technology problem. In A. Valdez, G. Scobie, & J.L. Dillon (Eds.), *Economics and the design of small-farmer technology* (pp. 167–177). Ames: Iowa State University Press.

Doob, L.W. (1960). *Becoming more civilized: A psychological exploration*. New Haven, CT: Yale University Press.

Freire, P. (1970). *Pedagogy of the oppressed*. New York: Herder & Herder.

Galbraith, J.K. (1979). *The nature of mass poverty*. Cambridge, MA: Harvard University Press.

Goodell, G.E. (1984). Untying the HYV package: A Filipino farmer grapples with the new technology. *Journal of Peasant Studies, 11*, 238–266.

Gore, P.S., & Rotter, J.B. (1963). A personality correlate of social action. *Journal of Personality, 31*, 58–64.

Gurin, P., & Katz, D. (1966). *Motivation and aspiration in the Negro college*. Ann Arbor, MI: University of Michigan Institute for Social Research.

Inkeles, A., & Smith, D. (1974). *Becoming modern*. Cambridge, MA: Harvard University Press.

Jorgenson, D.W. (1969). The role of agriculture in economic development: Classical vs. neoclassical models of growth. In C.R. Wharton (Ed.), *Subsistence agriculture and economic development*. Chicago: Aldine

Laosa, L.M. (1976). Developing arithmetic skills among rural villagers in Ecuador through nonformal education: A field experiment. *Journal of Educational Psychology, 68*, 670–679.

Lefcourt, H.M. (1976). *Locus of control: Current trends in theory and research*. Hillsdale, NJ: Erlbaum.

Lefcourt, H.M. (Ed.). (1981). *Research with the locus of control construct: Vol. 1. Assessment methods*. New York: Academic.

Lerner, D. (1958). *The passing of traditional society: Modernizing the Middle East.* Glencoe, IL: The Free Press.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H.M. Lefcourt (Ed.), *Research with the locus of control construct: Vol. 1. Assessment methods* (pp. 15–63). New York: Academic.

McClelland, D. (1977). The psychological causes and consequences of modernization: An Ethiopian case study. *Economic Development and Cultural Change, 25*, 43–66.

McClelland, D., & Winter, D. (1969). *Motivating economic achievement.* New York: The Free Press.

Mead, M. (1955). *Cultural patterns and technical change.* New York: New American Library.

Migdal, J.S. (1974). *Peasants, politics, and revolution.* Princeton, NJ: Princeton University Press.

Miller, S.M. (1980). Why having control reduces stress: If I can stop the roller coaster, I don't want to get off. In J. Garber & M. Seligman (Eds.), *Human helplessness* (pp.71–95). New York: Academic.

Moris, J. (1981). *Managing induced rural development.* Bloomington, IN: International Development Institute.

Oficina de Asesores del Presidente. (1980). *Primer planteamiento de metas de consumo y estrategia de producción de alimentos básicos para 1980–1982.* Mexico City: SAM.

Payne, M. & Furnham, A. (1985). Explaining the causes of poverty in the West Indies: A cross-cultural comparison. *Journal of Economic Psychology, 6*, 215-229.

Putterman, L. (1985). Extrinsic versus intrinsic problems of agricultural cooperation: Anti-incentivism in Tanzania and China. *Journal of Development Studies, 21*, 175–204.

Ramirez, M. (1983). *Psychology of the Americas: Mestizo perspectives on personality and mental health.* New York: Pergamon.

Rogers, E. (1962). *Diffusion of innovations.* New York: The Free Press.

Rotter, J.B. (1954). *Social learning and clinical psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Roueche, J.E., & Mink, O.G. (1976). *Impact of instruction and counseling on high risk youth.* Washington, DC: National Institute of Mental Health. (NIMH Grant R01MH22590)

Roumasset, J.A. (1977). *Risk and uncertainty in agricultural development* (Seminar Report No. 15). New York: Agricultural Development Council.

Sanders, J.H., & de Hollanda, A.D. (1979). Technology design for semiarid Northeast Brazil. In A. Valdez, G. Scobie, & J.L. Dillon (Eds.), *Economics and the design of small-farmer technology* (pp. 102–118). Ames: Iowa State University Press.

Scandizzo, P. (1979). Implications of sharecropping for technology design in Northeast Brazil. In A. Valdez, G. Scobie, & J.L. Dillon (Eds.), *Economics and the design of small farmer technology* (pp. 122–135). Ames: Iowa State University Press.

Schultz, T.W. (1964). *Transforming traditional agriculture*. New Haven. CT: Yale University Press.

Singh, S. (1983). Effect of motivation, values, cognitive factors, and child-rearing attitudes on productivity among Punjab farmers. *Journal of Social Psychology, 120*, 273–278.

Strickland, B.R. (1965). The prediction of social action from a dimension of internal–external control. *Journal of Social Psychology, 66*, 353–358.

Zandstra, H., Swanberg, K., Zulberti, C., & Nestel, B. (1979). *Caqueza: Living rural development*. Ottawa, ONT: International Development Research Centre.

Received December 16, 1986
Revision received July 24, 1987

AUTHOR

SHARON ALISA LOBEL. *Institutional affiliations and titles:* Universidade Gama Filho, Rio de Janeiro, Brazil; Professor. University of Michigan, Ann Arbor, Michigan, U.S.A.; Visiting Assistant Professor. *Degrees:* Ph.D., 1984, and M.A., 1981, Harvard University, U.S.A. *Specializations:* Cross-cultural research; organization behavior; human resource management. *Mailing address:* School of Business Administration, University of Michigan, Ann Arbor, MI 48109, U.S.A.

SUMÁRIO EXTENSO

Causas da Pobreza Rural Endêmica: Perspectivas Comportamentais

SHARON ALISA LOBEL

Diversas abordagens do ponto de vista comportamental têm sido apresentadas para explicar as causas da pobreza rural endêmica. Entre estas, as abordagens que apresentam o camponês como *tradicional*, como pessoa que *evita os riscos* , e como *desamparado* são as mais representativas.

A análise tradicional do camponês aponta aspectos de personalidade, tais como baixa necessidade de realização, que predispõe os camponeses a resistirem à adoção de inovações tecnológicas. As estratégias de desenvolvimento rural indicadas por esta abordagem incluem, primeiramente, a seleção de pessoas já predispostas à modernidade para receber a intervenção e, em segundo lugar, a orientação de camponeses para que eles adotem atitudes modernas. A limitação principal desta análise é que ela não leva em consideração o ambiente do camponês, o qual freqüentemente restringe rigidamente a escolha de um comportamento de mudança.

A abordagem que sugere que os camponeses evitam riscos deriva do campo da microeconomia e focaliza as avaliações racionais que os camponeses fazem do risco existente envolvido no processo de adoção de uma nova tecnologia. A estratégia indicada para desenvolvimento rural é a redução do risco por meio de inovações tecnológicas, por exemplo, sementes que resistem à seca. Os problemas associados com a definição e medição de risco ainda ficam para ser resolvidos. Especificamente, embora os fatores de risco associados com o meio social do camponês tenham sido comprovados significativos, eles não têm sido incorporados de uma forma adequada nos modelos do comportamento do camponês em tomar decisões.

A perspectiva que considera o camponês como sendo desamparado explica que os camponeses que percebem falta de controle pessoal sobre sua produção agrícola são menos capazes de desenvolver comportamentos que possam levá - los a melhorar a sua condição de vida. A estratégia de desenvolvimento rural indicada envolve a orientação para que os camponeses reconheçam os fatores internos e externos de controle que afetam a produção de uma maneira previsível. Já que as crenças de controle de um grupo específico derivam, em parte, das oportunidades existentes no ambiente, as intervenções têm que incluir avaliações realistas da razão pela qual os indivíduos endossam dadas crenças.

Após uma revisão e crítica de cada abordagem comportamental, é sugerido que a falta de atenção para os fatores ambientais têm enfraquecido a capacidade das perspectivas comportamentais em explicar e resolver o problema da pobreza crônica rural. No entanto, as abordagens puramente estruturais não garantem gerar estratégias que vão funcionar mais efetivamente porque, sobretudo, os programas têm que ser adaptados às condições locais. Além disso as variáveis comportamentais—tais como

os incentivos dos camponeses, o tipo de liderança na comunidade, a cooperação e a alienação em vários grupos, a motivação e capacidade do pessoal, e as atitudes com respeito ao programa e aos camponeses—todas têm um papel determinante no sucesso das intervenções. As perspectivas comportamentais, em geral, têm valor embora as abordagens utilizadas até agora tenham sido deficientes. Duas indicações para cientistas que têm interesse nos aspectos comportamentais do desenvolvimento rural são: (a) melhorias nas práticas de administração; e (b) utilização de pontos de vista locais. Recomendações específicas para melhor administração e estratégias de intervenção baseadas nos pontos de vista locais (Freire, Ramirez) são abordadas na conclusão.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp. 90-100

La Psicología Ambiental: Un Nuevo Campo de Aplicación de la Psicología y un Nuevo Rol Profesional Para el Psicólogo

EUCLIDES SÁNCHEZ

y

ESTHER WIESENFELD

Universidad Central de Venezuela
VENEZUELA

La psicología ambiental tiene como propósito el aporte de información psicológica para la comprensión de la interacción entre las personas y el ambiente. Este objeto de estudio demanda del psicólogo que actúa en el área de ambiente y conducta, características profesionales diferentes a las exigencias tradicionales del ejercicio profesional de la psicología. El análisis de este nuevo campo de aplicación de la psicología, en cuanto a sus implicaciones para la práctica del psicólogo, constituye el objetivo de este artículo, para lo cual revisaremos el desarrollo de la dsiciplina, describiremos un marco de referencia para la práctica del psicólogo ambiental y propondremos algunas directrices que pueden orientar la capacitación en los roles que se señalan.

La psicología ambiental, disciplina de reciente aparición en las ciencias del comportamiento, nace como respuesta a la necesidad de completar, desde el punto de vista psicológico, la explicación de la interrelación entre las personas con su ambiente, entendido éste como totalidad. Este objeto de estudio, postergado en la psicología por muchos años, le impone características especiales a la práctica del psicólogo en el área de ambiente y conducta, cuyo análisis constituye el propósito de éste trabajo. Con este fin examinaremos los antecedentes, características y tópicos de estudio de este nuevo campo de aplicación de la psicología y discutiremos los tipos de actividades y exigencias de formación inherentes al ejercicio profesional en esta área.

Copyright © 1987 by Euclides Sanchez and Esther Wiesenfeld. All rights reserved.

ANTECEDENTES

La psicología ambiental surge bajo la influencia de condiciones sociales e intelectuales entre las cuales se destacan: (a) El desarrollo de un conjunto de investigaciones pioneras que proveyeron datos sobre aspectos específicos relativos a la conducta y al ambiente, tales como la importancia del espacio personal en la interacción social (Hall, 1966), la viabilidad metodológica del abordaje sistemático de la interrelación lugar-organismo-conducta (Barker, 1968), el impacto del diseño interior en la comunicación interpersonal (Sommer, 1974). (b) El reconocimiento por parte de disciplinas científicas, tanto naturales como sociales, del impacto de problemas como la alta densidad, la contaminación y el déficit energético, sobre el bienestar humano. (c) La demanda a los psicólogos por parte de sociólogos, diseñadores y arquitectos, de explicaciones sobre el papel de variables psicológicas (como las necesidades, preferencias, actitudes, percepciones) en el comportamiento de las personas en el ambiente construído. (d) La toma de conciencia de los profesionales de la psicología acerca de la ausencia de conocimientos relevantes en la psicología que permitiesen dar respuestas a las peticiones provenientes de otras disciplinas, hecho que estimuló en la psicología la apertura de líneas de investigación, dirigidas a obtener información sobre la relación entre variables psicológicas y ambientales y a desarrollar metodologías apropiadas al estudio de la interacción entre los individuos y su entorno. (e) La creación de organizaciones (Environmental Design Research Association, International Association for the Study of Man and his Physical Surrounding, la División de Psicología Ambiental y Población de la American Psychological Association), de revistas científicas (*Environment and Behavior, Journal of Environmental Psychology, Population: Behavioral, Social and Environmental Issues*) y de cursos de psicología ambiental tanto a nivel de pre-grado como de post-grado que han fomentado la divulgación de la disciplina en sectores que constituyen usuarios potenciales de su conocimiento (diseñadores, planificadores urbanos, sociólogos), la formación de recursos humanos con profesionales provenientes de diferentes disciplinas (arquitectura, sociología) y el desarrollo teórico, metodológico y aplicado de la psicología ambiental.

DEFINICIÓN

La definición de psicología ambiental se ha ido enriqueciendo con el desarrollo mismo de la disciplina. Así, en una de las primeras definiciones, la de Proshansky (1970) se considera que la psicología ambiental "es el estudio de la conducta humana en relación con el ambiente ordenado y definido por el hombre" (p. 5). En 1977, Stokols la concibe como "el estudio de procesos psicológicos básicos como la percepción, la cognición, el desarrollo de la personalidad y el aprendizaje social, como mediadores

de la relación entre la conducta humana y el ambiente" (p. 12). Más recientemente Canter y Craik (1981) la definen como "el área de la psicología que integra y analiza las transacciones e interrelaciones de las experiencias y acciones humanas con aspectos pertinentes de los entornos socio-físicos" (p. 2). Comparando estas tres definiciones notamos el reconocimiento cada vez mayor por parte de los autores del papel mediador de los procesos psicológicos en la relación persona–ambiente y de la concepción interaccionista de este relación. Esta concepción interaccionista de la conducta, aunque ya había sido formulada por Lewin en 1936, no es sino hasta fines de la década del 60 cuando un mayor número de psicólogos sociales se interesa por desarrollarla.

CARACTERÍSTICAS

Las diversas revisiones que se han hecho acerca de la psicología ambiental coinciden en atribuirle una serie de características propias que en general la diferencian de la psicología tradicional. Así, la psicología tradicionalmente ha estudiado al individuo separándolo de su contexto y ha considerado el ambiente como pequeñas unidades moleculares aisladas tales como estímulos visuales y auditivos; ha asumido un modelo determinista de la conducta, según el cual ésta es concebida como la reacción pasiva a las condiciones ambientales; ha utilizado un enfoque predominantemente intradisciplinario, en el estudio y explicación del comportamiento, estando más motivada al desarrollo de conocimientos teóricos que a su aplicación. En contraposición, la psicología ambiental estudia a los seres humanos en el contexto en el cual ocurren sus acciones, preservando la relación integral existente entre la conducta y el entorno físico en que ella se realiza (Proshansky, 1970) o, lo que es lo mismo, adoptando un enfoque holístico en el estudio de dicha relación (Holahan, 1982). Uno de los ejemplos que ilustra esta característica es el estudio de Saegert, Mackintosh y West (1975) en el cual se investigó la influencia de la aglomeración en los procesos cognitivos y afectivos. Con este fin los autores expusieron una muestra de sujetos de ambos sexos a condiciones de baja y alta densidad, en una estación de ferrocarril en la cual el flujo natural de usuarios en diferentes períodos del día determinaba las condiciones de densidad seleccionadas para la investigación. Los resultados demostraron que el aumento de aglomeración disminuyó la ejecución de tareas encomendadas a los sujetos e incrementó emociones negativas como la ansiedad y la agresividad. Como vemos, el estudio se realizó respetando la relación natural que ocurría entre los sujetos y las condiciones de la estación, relación que variaba según la dinámica de la influencia mutua entre las condiciones ambientales y los estados cognitivos y afectivos de las personas expuestas a las condiciones de densidad. Esta perspectiva ecológica de la psicología ambiental implica además que la relación entre la persona y su ambiente es concebida como un intercambio dinámico en el cual el individuo cambia el ambiente y estos cambios a su vez influyen sobre su conducta.

En esta concepción, la persona es vista como un ente activo cuyos procesos psicológicos intervienen en la generación de alternativas de acción que puedan modificar las condiciones ambientales a fin de que éstas se ajusten a las expectativas y necesidades de los individuos. Una ilustración de esto es el papel que tienen algunas variables de personalidad en la percepción de desastres naturales. Así, por ejemplo, Sims y Bauman (1972) encontraron que el número de víctimas que los tornados producen en el sur y medio oeste de los Estados Unidos tiene que ver con la percepción de peligro de los residentes y con las correspondientes previsiones que se adoptan. Mientras que en el oeste se producen más tormentas que en el sur, el número de muertes es mayor en el sur. Los autores explican este fenómeno en base a las diferencias de locus de control encontradas entre las personas del oeste y del sur. Las primeras tienden a creer más en sus propias acciones, en el sentido de que la adopción de medidas de seguridad contribuirá a reducir las consecuencias de los ciclones, mientras que las segundas, al asumir que el azar o el destino es el responsable de lo que ocurra, no despliegan comportamientos que ayuden a reducir las consecuencias de estos fenómenos naturales.

Por otro lado, la relación dinámica y compleja que se da entre conductas, procesos psicológicos y ambiente, demanda explicaciones que trascienden lo meramente psicológico. En efecto, por una parte la psicología ambiental es un campo del cual se nutren profesionales de disciplinas como la sociología, la geografía, la antropología, la arquitectura, la planificación, así como también la psicología clínica, industrial, organizacional, educativa, social. Sus hallazgos empíricos y sus fundamentos teóricos añaden una dimensión más a la comprensión de la conducta humana en entornos físicos. De la misma manera, la psicología ambiental no puede ignorar los factores arquitectónicos, climáticos, sociales y culturales, implícitos en la interacción humana–ambiental. De allí que la psicología ambiental debe ser vista como otra área de las ciencias ambientales y sociales y no como disciplina aislada, lo cual le imprime un carácter interdisciplinario.

La psicología ambiental, adicionalmente, está orientada tanto a lo teórico como a lo aplicado, por lo cual la investigación que se realiza en ella se propone tanto el desarrollo de teorías como su empleo para optimizar el intercambio hombre–ambiente. El conocimiento se valora tanto por su relevancia teórica como por su utilidad. Esta doble finalidad de la investigación, ya formulada por Lewin en 1936, pero puesta en práctica por la psicología muy recientemente, permite incrementar el conocimiento científico y tecnológico de la disciplina. Un caso que ejemplifica estas dos últimas características es el proyecto Casalta, un proyecto de autoconstrucción de viviendas en el cual hemos estado involucrados. En esta oportunidad los psicólogos trabajaron con arquitectos a fin de conocer mediante simulación la percepción de los usuarios acerca de sus viviendas, de modo que esta información orientase las decisiones del diseño. Por otro lado la participación de las personas tanto en el diseño como en la construcción exigía un modelo de organización de la comunidad flexible, adaptable a los requerimientos de las diversas tareas y democrático, que facilitase la incorporación activa de los miembros de la comunidad en la toma de decisiones. La colaboración de los

psicólogos en el logro de este objetivo requirió la realización de investigaciones cuyos resultados debían ser aplicados inmediatamente y el uso de conocimientos provenientes de disciplinas como la sociología (sociología de los movimientos urbanos), psicología social (actitudes, roles, creencias) y técnicas de dinámica de grupo.

TÓPICOS DE ESTUDIO

Los temas de estudio en la psicología ambiental han sido clasificados por Stokols (1977), según el tipo de transacción que se da entre las personas y su medio ambiente. Stokols identifica los siguientes modos de transacción: *el interpretativo* que se refiere a la manera en la que el individuo conoce y se orienta en el ambiente, *el operativo* que engloba las acciones sobre dicho ambiente, *el evaluativo* que apunta a los juicios acerca de la calidad de los diferentes entornos y *el respondiente* que agrupa el impacto psicológico de factores ambientales.

En el modo interpretativo se incluye el estudio de los procesos perceptuales, cognitivos y afectivos y rasgos de personalidad relacionados con el coocimiento y formas de transacciones con el ambiente. En el operativo se estudian las consecuencias del comportamiento humano, tales como conductas deteriorantes sobre el ambiente (contaminación, basura) y el uso de mecanismos como la privacidad, territorialidad y el espacio personal, para regular la interacción. En el evaluativo se exploran las preferencias por futuros entornos y la satisfacción con edificaciones existentes. En el repondiente interesa conocer el impacto de estresores ambientales (ruido, alta densidad, calor) y del ambiente natural (clima) y construido (diseño urbano, vivienda).

MARCO DE REFERENCIA PARA LA PRÁCTICA DEL PSICÓLOGO AMBIENTAL

La actuación profesional del psicólogo ambiental impone que dicha práctica esté orientada por una concepción que la haga congruente con los requerimientos de la disciplina, esto es, que refleje la interdependencia existente entre la gente y su ambiente, el carácter dinámico de esta relación y el papel activo que juegan los individuos en la misma; que utilice enfoques provenientes de disciplinas que tienen que ver con diferentes dimensiones de la relación ambiente–comportamiento y que vincule la producción de conocimientos teóricos con la solución de problemas.

A nuestro juicio dicha concepción puede inspirarse en algunas de las proposiciones formuladas por autores interesados en promover una psicología social aplicada, tal como la de Varela (1971) que, al igual que en la psicología ambiental, tenga un desarrollo científico orientado tanto a la construcción del conocimiento teórico como a su utilización a través de intervenciones para la resolución de problemas socialmente relevantes. En este sentido Sánchez y Wiesenfeld (1982), luego de un análisis de la

crisis de la psicología social y de las proposiciones formuladas para superarla, sugieren la integración de la investigación básica, la investigación aplicada y la intervención tecnológica, a fin de superar la tradicional separación entre lo teórico y lo aplicado. La implementación de esta proposición requiere la participación directa de los actores vinculados al problema en las diversas etapas del proceso. Dichas etapas, basadas en los planteamientos de Reyes y Varela (1980), son: (a) Detección del problema, evaluación de su importancia social y de la factibilidad de ser abordado. (b) Diagnóstico del problema que contempla la recolección de datos a fin de precisar sus determinantes y conceptualizarlas en términos de constructos teóricos. (c) Selección de los cursos de acción a seguir en función del estado de los conocimientos existentes sobre el problema. Es decir, si no existe un mínimo de información teórica y/o aplicada que oriente la intervención, se requiere la planificación de investigaciones que provean dicho conocimiento. Si éste no es el caso, se procede a diseñar la intervención para la solución del problema, transformando los conocimientos existentes en instrumento de acción. (d) Ejecución de la intervención. (e) Evaluación de la acción realizada en términos de su impacto sobre el problema y de la eficiencia de sus apoyos teóricos a fin de determinar la necesidad de nuevas acciones sobre el problema y de estimar el valor práctico (utilidad) de la intervención en problemas similares.

Esta propuesta trasladada al área de ambiente y conducta exige que la investigación y la planificación ambiental sean consideradas como actividades relacionadas dentro de un proceso multifásico. Los roles que el psicólogo ambiental desempeñe estarán vinculados en consecuencia a los del planificador ambiental y se orientarán por el conocimiento teórico existente y por el modelo de la participación en su acción teórica y práctica. De este modo se garantizaría lo que Stokols (1977) ha denominado la vitalidad futura de la psicología ambiental, según la cual la existencia de una relación equilibrada entre la producción de teorías y su aplicación evitaría que la disciplina se convirtiese en "ciencia estéril" o "disciplina empírica".

Por otra parte, el desarrollo de un cuerpo teórico propio en la psicología ambiental, que refleje la naturaleza de las transacciones conducta–ambiente, es particularmente necesario, ya que debido a la juventud de la disciplina gran parte de sus constructos teóricos son prestados de otras disciplinas. Estos constructos enfatizan el papel de los procesos interpersonales que median el impacto del ambiente sobre la conducta y no el carácter contextual de las relaciones entre las personas y su entorno. Los constructos contextuales, por el contrario, se refieren a los procesos y eventos que dentro del ambiente socio-físico afectan la experiencia y conducta de individuos y grupos, así como la acción transformadora de éstas sobre las condiciones ambientales (Stokols, 1977). Este desarrollo teórico permitiría además sistematizar el agregado de datos empíricos dispersos, proveer explicaciones a los hallazgos contradictorios generados dentro de la psicología ambiental, llenar algunas lagunas en el conocimiento e identificar áreas que requieren nuevos estudios (Rapoport, 1973).

Desde otro punto de vista, al existir una vinculación permanente entre la investigación y sus hallazgos para la solución de problemas ambientales, se solucionaría el problema del uso infrecuente de la teoría en el abordaje práctico de cuestiones

ambientales, es decir, de la aplicación del conocimiento científico en la planificación y evaluación ambiental.

Adicionalmente, dado que un aspecto central en el enfoque propuesto lo constituye la participación de las personas en las diferentes etapas de abordaje del problema, se estaría facilitando un papel activo de aquéllas para el control personal del ambiente o la realización del *proyecto personal* que, según Little (1983), consiste en el conjunto de actividades realizadas en entornos y momentos determinados, dirigidas al logro de metas y planes importantes.

Roles en la Psicología Ambiental

Los diferentes roles que el psicólogo ambiental desempeña pueden agruparse de acuerdo a la clasificación propuesta por Hornstein (1975) en relación con los roles del psicólogo social en: investigador, orientador y tecnólogo social. Como investigador el psicólogo ambiental puede desarrollar investigaciones básicas, con el propósito de obtener conocimiento acerca de las relaciones entre variables intrínsecas a los problemas sociales estudiados, como lo es la relación entre densidad y estrés e investigaciones aplicadas, que proporcionen información precisa que oriente la solución inmediata de los problemas. Este tipo de investigación puede ser de carácter diagnóstico, dirigida a caracterizar la situación problemática en una organización, grupo o comunidad, a fin de orientar las acciones requeridas para su intervención, por ejemplo, las actitudes hacia medidas que requieren colaboración ciudadana para controlar la contaminación; investigación evaluativa destinada a estimar el impacto de programas ambientales como por ejemplo los de desarrollo habitacional y las campañas de conservación de energía y de conservación de las aguas.

En el rol de orientador, el psicólogo ambiental tiene la función de comunicar información científica que contribuya a cambiar las condiciones ambientales existentes. Esta comunicación puede afectar el diagnóstico de los problemas o su solución. Por ejemplo, el psicólogo puede explicar a un grupo de planificadores urbanos y usuarios, los resultados de las investigaciones sobre los efectos de la densidad en el comportamiento, a objeto de influir en las decisiones del plan urbano. Este logro puede darse con colaboración institucional o en ausencia de ella, trabajando directamente con los grupos afectados por los problemas ambientales.

El rol de tecnólogo social puede describirse según la concepción que se adopte en relación con la tecnología social. Así, de acuerdo con lo postulado por Varela (1975) y Reyes y Varela (1980), el tecnólogo social utiliza conocimientos disponibles en la psicología para diseñar estrategias de intervención dirigidas a cambiar los comportamientos de las personas y así favorecer su adaptación al ambiente. Por otra parte, para Wiesenfeld y Sánchez (1983) y Sánchez y Wiesenfeld (1983) el tecnólogo social diseña estrategias de intervención basándose en conocimientos existentes, con la finalidad de estimular la participación de las personas en la modificación de las condiciones ambientales. En esta concepción, se enfatiza la necesidad de evaluar la intervención a fin de medir los efectos de ésta y de conocer sus implicaciones para el

desarrollo de la metodología y de los principios teóricos que sustentaron la tecnología empleada. Un ejemplo que ilustra este enfoque es el trabajo conjunto de psicólogos ambientales, ingenieros sanitarios, arquitectos, educadores y usuarios para identificar los determinantes del deterioro ambiental en una ciudad universitaria e implementar acciones correctivas en las cuales el usuario sea el agente principal en la restauración de dicho entorno.

De acuerdo con nuestros planteamientos iniciales, sobre los postulados que orientan estos roles, es de suponer que las funciones descritas tienden a estar vinculadas una con la otra en ese proceso que hemos denominado multifásico.

La Formación del Psicólogo Ambiental

En esta sección no pretendemos proponer un plan curricular detallado para la formación de recursos humanos en el área de ambiente y conducta, sino un esquema general que establezca ciertas directrices para el entrenamiento en los roles definidos.

El modelo de formación del especialista en psicología ambiental debe diseñarse en función del logro de objetivos compatibles con el carácter de la disciplina. Así, las diversas actividades en enseñanza deben tener como metas: (a) Concientizar acerca de la vinculación natural entre el comportamiento y el ambiente; acerca de la relación ecológica, sistemática entre ambos, lo cual se contrapone a la noción de conducta autónoma. (b) Hacer entender el papel activo con que la persona se enfrenta a las condiciones ambientales. No es un procesador de información unicamente quien responde, sino un agente que modifica su propio entorno. (c) Desarrollar actitudes y destrezas para el trabajo interdisciplinario. El trabajo en equipo entre diferentes profesiones no es lo usual, por el contrario la enseñanza universitaria propicia la separación, el desconocimiento de los aportes de cada disciplina, de allí la importancia de estimular el reconocimiento de los enfoques y metodologías de otras áreas. Pero no es suficiente valorar positivamente el trabajo conjunto; se deben desarrollar habilidades que permitan resolver los diferentes obstáculos que surgen al unirse diversos enfoques de los problemas. (d) Fomentar la producción del conocimiento y su transformación en cursos de acción para solucionar los problemas inherentes a la interacción del ser humano con su ambiente.

Para lograr estos fines la estructura curricular del plan de formación debe estar organizada en base a tres dominios. En primer lugar, un conjunto de conocimientos sustantivos en áreas de diseño y planificación ambiental, psicología, particularmente psicología social y ambiental y otras ciencias sociales, (sociología, antropología, economía, politología, geografía). Segundo, conocimientos metodológicos en estrategia y técnicas de investigación y en procedimientos de análisis estadísticos. Tercero, actividades de trabajo de campo (realización de investigaciones, pasantías) mediante las cuales se fomente la participación de los usuarios y de profesionales de varias disciplinas y el contacto con problemas de distinta naturaleza, que incluyan tanto la satisfacción de la persona y del grupo con microambientes como la vivienda, el lugar

de trabajo, la escuela; macroambientes como el vecindario, la ciudad; estresores ambientales como contaminación, densidad, ruido; la escasez de recursos energéticos; los incendios forestales; la organización de las comunidades para el mejoramiento de su habitat.

Finalmente el contacto con los problemas mencionados debe contemplar tanto la descripción del problema en estudio como su intento de solución con la participación de las personas y el uso de los conocimientos derivados de las disciplinas relevantes al problema.

NOTA

Un análisis exhaustivo del origen, desarrollo e importancia de la psicología ambiental en América Latina se encuentra en E. Sánchez, E. Wiesenfeld & K. Cronick (1987), Environmental psychology from a Latin American perspective, en D. Stokols e I. Altman (Eds.), *Handbook of environmental psychology*. New York: Wiley.

REFERENCIAS

Barker, R. (1968). *Ecological psychology*. Stanford: Stanford University Press.

Canter, D. & Craik, K. (1981). Environmental psychology. *Journal of Environmental Psychology, 1*(1), 1–12.

Hall, E. (1966). *The hidden dimension*. New York: Doubleday.

Holahan, C. (1982). *Environmental psychology*. New York: Random House.

Hornstein, H. (1975). Social intervention. En M. Deutsch & M. Hornstein (Eds.), *Applying social psychology* (pp. 211–234). New York: Wiley.

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Little, B.K. (1983). Personal projects: A rationale and method for investigation. *Environment and Behavior, 15*(3), 273–310.

Proshansky, H. (1970). Introduction. En H. Proshansky, W. Ittelson & L. Rivlin (Eds.), *Environmental psychology* (pp. 1–6. New York: Holt, Rinehart and Winston.

Rapoport, A. (1973). An approach to the construction of man–environment theory. En W. Preiser (Ed.), *Proceedings of the 4th Environmental Design Research Association Conference* (Vol. 2, pp. 124-135). Stroudsburg, Pennsylvania: Dowden, Hutchinson & Ross.

Reyes, H., & Varela, J. (1980). Conditions required for a technology of social sciences. En R. Kidd & M.J. Saks (Eds.), *Advances in applied social psychology* (Vol. 1, pp. 47–80). New Jersey: Erlbaum.

Saegert, S., Mackintosh, E., & West, S. (1975). Two studies of crowding in urban public places. *Environment and Behavior, 7*(2), 159–184.

Sánchez, E., & Wiesenfeld, E. (1982). La psicología social aplicada: Algunas consideraciones. *Boletín de AVEPSO, 5*(3), 7–14.

Sánchez, E., & Wiesenfeld, E. (1983). Psicología social aplicada y participación: Metodología general. *Boletín de AVEPSO, 6*(3), 20–27.

Sims, J.M., & Bauman, D.D. (1972). The tornado threat: Coping styles of the north and the south. *Science, 176*, 1386–1391.

Sommer, R. (1974). *Espacio y comportamiento individual*. Madrid: Nuevo Urbanismo.

Stokols, D. (1977). Origins and directions of environment–behavior research. En D. Stokols (Ed.), *Perspectives on environment and behavior* (pp. 5–36). New York: Plenum.

Varela, J. (1971). *Psychological solutions to social problems: An introduction to social technology*. New York: Academic.

Varela, J. (1975). Can social psychology be applied? En M. Deutsch & H. Hornstein (Eds.), *opus cit.* 157–174.

Wiesenfeld, E. & Sánchez, E. (1983). La tecnología social: Análisis de algunos casos. *Boletín de AVEPSO, 6*(1), 3–10.

Recibido en abril 22 de 1986
Revisión recibida en diciembre 15 de 1986

AUTORES

EUCLIDES SÁNCHEZ. *Afiliación institucional:* Universidad Central de Venezuela, Caracas, Venezuela. *Título:* Jefe de Departamento (Psicólogo). *Diploma:* Master of Science, 1973, London School of Economics, London, England. *Especializaciones:* Psicología social y ambiental. *Dirección de correo:* Apartado 47018, Caracas, 1041-A, Venezuela.

ESTHER WIESENFELD. *Afiliación institucional:* Universidad Central de Venezuela, Caracas, Venezuela. *Título:* Investigador, Docente (Psicólogo). *Diploma:* Magister en Psicología Social, 1984, Universidad Central de Venezuela. *Especializaciones:* Psicología social y ambiental. *Dirección de correo:* Apartado 47018, Caracas, 1041-A, Venezuela.

EXTENDED SUMMARY

Environmental Psychology: A New Field of Application in Psychology and a New Professional Role for the Psychologist

EUCLIDES SÁNCHEZ

and

ESTHER WIESENFELD

Environmental psychology is a discipline that combines psychology and the environmental sciences. This combination falls within a conceptual structure that frames the understanding of person–environment relationships in terms that emphasize the interdependence of both parameters. Different definitions have been given of the discipline, and these have become increasingly complex in their evolution so as now to include specific human and environmental components. For example, in the earliest definitions no reference was made to the intervening psychological processes, the acknowledgment of which is now considered obligatory. The evolution of the discipline can be characterized by the topics that are now studied, the academic conceptions that now guide the formation of the professionals in the field, and the assumptions now made by the practitioners in the process of applying the knowledge that has been obtained.

It is not possible to consider human behavior in environmental psychology without linking it to the surroundings where the behavior takes place. The influence of the surroundings can be greater or lesser in scale, but this variability does not make its effects on the personality less relevant. Since an integral relation between the person and the environment does exist, behavior has to be examined not only for its background but also for the individual's active participation in the decisions that affect him or her. This posture requires that theory and practice be linked in an interactive conceptualization so that the person's control of environmental transformations may be stimulated.

The discussion of these issues is the purpose of this paper. We examine the antecedents, the definitions, the characteristics, and the topics of study in environmental psychology, and we propose a frame of reference for the practice of the environmental psychologist.

Revista Interamericana de Psicología/Interamerican Journal of Psychology
1987, Vol. 21, Nos. 1 & 2, pp. 101-107

Libros/Books

THEORIZING ABOUT THEORIES

The Nature of Psychological Explanation, by Robert Cummins. Cambridge, Massachusetts, USA: The MIT Press, 1983. x + 219 pp. US$22.50.

Cummins believes that psychological explanation has not been properly analyzed by philosophers. For this reason, the book under review is about the nature of psychological explanation:

> Psychologists theorize about animals—humans, mainly—in an effort to understand their dispositions and abilities. When the theorizing is good theorizing, it helps us to understand such things as your ability to calculate, speak a language, remember your name and way home, and tell your elbow from a hot rock. Good psychological theory explains such things (or tries to) . . . I theorize about this phenomenon, i.e., about the fact that psychological theory can make us understand such abilities as perception and cognition. How does it do this? What would it take to make us understand the capacity to be operantly conditioned or the capacity to learn French? (p. v)

In Cummins' view, to understand is to analyze. He admits that there is nothing revolutionary about this idea, and that much explanation in psychology is by analysis. Cummins' particular concern, however, is with the appropriate *kind* of explanation required for cognitive capacities. Specifically, he shows that adherence to the deductive nomological model of the positivists—a system that has dominated the field at least since the time of Newton—has in fact distorted the view of philosophers of science about what psychologists do and thereby masked the real nature of explanation. On the other hand, he seemingly has little concern for the truth of a particular theory and reviews the systems of Titchener, Hull, and Freud only to show that "most psychological explanation makes little sense construed as causal subsumption" (p. 118).

Nineteen discrete sections in the text are organized into four chapters: Analysis and Subsumption (4 sections), Functional Analysis (3), Understanding Cognitive Capacities (9), and Historical Reflections (3). Understanding Cognitive Capacities, the chapter on the major issues involved in explaining cognitive capacity in humans, probably has the most relevance for psychologists and is also given the most attention (66 pages) by Cummins. Approximately one-fourth of the book is devoted to an appendix, notes, and references; a three-page combined author/subject index completes the work.

Robert Charles Cummins received his Ph.D. from the University of Michigan in 1970. He has taught at Johns Hopkins University, the University of Wisconsin, Milwaukee, and currently is associate professor of philosophy at the University of Illinois, Chicago Circle, where he has been since 1982.

Not written for the casual reader, *The Nature of Psychological Explanation* is an important contribution to the philosophy of science. It deals with a timely, albeit difficult, topic and should be of interest to scholars in several disciplines and to advanced students in a variety of courses, including ones in artificial intelligence.

Jon D. Swartz

BOOK REVIEWER

JON D. SWARTZ. *Institutional affiliation:* Southwestern University, Georgetown, Texas, U.S.A. *Titles:* Associate Dean for Libraries and Learning Resources; Professor of Education and Psychology. *Degrees:* Ph.D., 1969, and Senior Post-Doctoral Fellowship in Community Psychology and Community Mental Health, 1973-74, The University of Texas at Austin, U.S.A. *Specialization:* Educational psychology. *Mailing address:* A. Frank Smith, Jr. Library Center, Southwestern University, Georgetown, TX 78626, U.S.A.

A REVIEW OF RESEARCH IN SPECIAL EDUCATION

Special Education: Research and Trends, by Richard J. Morris and Burton Blatt, editors. New York: Pergamon Press, 1986. vii + 419 pp. US$19.50 (paper).

The intended readers for this book are not specified, but in view of the research and trends focus it would appear most useful to graduate-level students, research workers, and professionals.

Twenty-three authors in twelve chapters attempt to review critically the research in special education and related fields over the past several decades and to isolate trends and problems pertinent to their particular specialty. The book also purports to cover "the entire range of handicapping conditions. . . ." (back cover). The authors are all well-known and well-regarded in their several specialties.

Chapter One presents an overview of special education research and a summary of the chapters to follow. Chapter Two (Ysseldyke) covers assessment in special education and its uses in decision making. Included is a brief historical overview of assessment, and the remaining content fairly represents the author's belief that special education should shift from discipline research to policy research—a position some readers might question. Chapter Three (Bricker) discusses the issues involved in early intervention programs and, through arguments pro and con on the issue, covers the literature from historical perspectives. Chapter Four (Morris and McReynolds) presents an overview of research covering a wide variety of behaviors that have been treated by behavior modification procedures, from assertiveness to yelling. The inclusion of tables in this lengthy, but well-written, chapter is quite helpful.

While research in special education has, as pointed out by Morris and Blatt, progressed, improved, and increased dramatically over the past two decades, this reviewer was left wondering about the effects of research on actual practice. Much of the text is consumed with issues such as definitions of various handicaps. Another common theme is equivocation (e.g., learning disabilities research). Offsetting this is Smith and Robinson's mention of the "lack of specificity in descriptions of the LD subjects used in research" (p. 237). The same caveat could be addressed to language research, emotional disturbance, sensory impairments, the gifted, and the severely handicapped. Relative to the chapter on the severely handicapped, for example, we learn that this handicapping condition encompasses "approximately the lowest intellectually functioning 1% of the school age population" and includes those labelled *"psychotic, autistic, moderately/severely/profoundly retarded, trainable level retarded, physically handicapped, multihandicapped, and deaf-blind"* (p. 131). On page 132, an IQ of less than 55 is mentioned as one aspect of a severe handicap. Little wonder that replication of research on the handicapped is rendered difficult! Such lack of specificity in describing research subjects echoes the warning of Smith and Robinson. Also, readers are treated to a somewhat overly optimistic portrayal of the "educability-

trainability" of those who are labelled profoundly mentally retarded a la AAMD classifications. Conversely, the moderately retarded or trainable, being lumped in the category of severely handicapped, gives the reader the impression that "moderates" are as severely impaired as the profoundly retarded. One can only be awestruck over the great breadth and scope of the "severely handicapped." Overlap from one developmental disability to others is inevitable; for example, profound retardation and/or hearing impaired often share speech and language handicaps, but the primary classification in profound retardation ordinarily is not too difficult to make: profound retardation with superimposed speech/language handicap.

Chapter Ten, dealing with sensory impairments, limits *sensory impairment* to hearing and vision (p. 303). The chemical senses of gustation and olfaction (as well as the sense of touch) are omitted. Such an omission testifies to the research lacuna that these "lesser" senses have enjoyed. The authors, Terzieff and Antia, do cover the two major senses well and also discuss technological developments such as the Optacon, hearing aids, captioned films, telecommunication devices for the deaf, laser canes, etc. Largely the products of engineers, these technical advances have greatly accelerated communicative and information processing among the hearing and visually impaired members of society.

The final chapter (Blake and Williams) addresses special education research in perspective. Here readers are reacquainted with individual differences via the new interest in qualitative research, and difficulties in conducting research are detailed.

In summary, how does this book compare with other texts in special education? Although it could be recommended for certain audiences (viz., those already mentioned), the latest revision of Kirk and Gallagher's introductory textbook hits a broader audience. Overall, *Special Education: Research and Trends* does an adequate job of reviewing research in the areas covered. Referencing is up-to-date and most chapters are well-written; but the claim of covering "the entire range of handicapping conditions" is not justified. Such an encyclopedic coverage still awaits publication.

Charles C. Cleland

BOOK REVIEWER

CHARLES C. CLELAND. *Institutional affiliation:* University of Texas at Austin, U.S.A. *Title:* Professor. *Degrees:* Ph.D., 1957, University of Texas at Austin. *Specializations:* Research with profoundly mentally retarded, organizational research, research on the antecedents of creativity. *Mailing address:* 3427 Monte Vista, Austin, TX 78731, U.S.A.

AMERICAN PSYCHOLOGY

Psychology in America: A Historical Survey, by Ernest R. Hilgard. Harcourt Brace Jovanovich, 1987. xxix + 1009 pp. US$34.95.

Hilgard has been a participant and an observer in American psychology for nearly sixty years, and he brings his vast experience to bear in this large survey of the development of psychology from its inception to the present time. Although some familiarity with psychology is assumed (the book is primarily for psychologists and psychology students) no specialized knowledge is required for understanding even the most difficult chapters. The relationship of a given area of research to the larger concerns of psychology or of science in general is stated plainly and persuasively. Technical topics are treated in sufficient detail to make apparent their significance, but not so thoroughly as to be intimidating or boring to the non-technical reader.

The organization of the book is primarily by topic area, although the first three chapters deal with particular historical periods; together they constitute an overview of the development of the profession prior to 1935. The next seventeen chapters trace the development of various subspecialties within the discipline and the profession. Chapter titles do not correspond directly to any particular schema of "schools" of psychology nor to the organization of the American Psychological Association: they reflect Hilgard's perception of the way that the topic may be presented most effectively. This organization will seem intuitively quite reasonable to most readers trained in psychology, perhaps because many of the chapter headings resemble those in Hilgard's excellent and extremely well known introductory psychology text. In addition to these familiar topics, three full chapters are devoted to what might be thought of as practitioner psychology (clinical, educational, and industrial/organizational psychology) and another to the professional organization of psychologists. The last chapter treats the question of psychology as a science and lists some of the trends that have developed during the history of psychology and some of the conflicts that have survived to the present.

The book is not nearly so fearsome as the page count would lead one to expect. There are just over 800 pages of text, enlivened by frequent brief biographical sketches of important participants, complete with photographs, and a reasonable sprinkling of other photographs, diagrams, charts, and graphs. Some of these are quite charming, as in the detailed diagrams of Wundt's Leipzig Laboratory and of the Stanford Psychology Laboratory of 1906. There are over 100 pages of references and another 43 pages of chapter notes, but the text flows very easily; reference citations within the text are not excessive, and footnotes are more often fascinating than intrusive. The only conceivable quarrels one might have with the book are its organization, which is surely as good as any other, and its size, which somehow seems less significant once one

begins to read. It is both an excellent text and an excellent resource document. It is well written, well documented, and well presented. If you like psychology, you will like *Psychology in America*. If you like both psychology and history, you will love it.

Robert C. Reinehr

BOOK REVIEWER

ROBERT C. REINEHR. *Institutional affiliation:* Southwestern University, Georgetown, Texas, U.S.A. *Title:* Assistant Professor of Psychology. *Degree:* Ph.D., 1965, University of Texas, U.S.A. *Specialization:* Clinical psychology. *Mailing address:* Box 6379, Southwestern University, Georgetown, TX 78626, U.S.A.

LAS DISFUNCIONES DEL LENGUAJE

Problemas de Pronunciación en el Niño: Evaluación Funcional y Programas Correctivos (Segunda Edición), por Guido Aguilar. Guatemala: Editorial Piedra Santa, 1987. 88 pp.

Este psicólogo centroamericano, autor de varias obras y artículos, nuevamente se ocupa de un tema relacionado con dificultades en el habla: los problemas de pronunciación en el niño. Errores en la articulación como el rotacismo (dificultad para pronunciar la rr), el sigmatismo o ceceo (s), mitacismo (m) etc. son tratados en este libro bajo tres aspectos generales: el estructural (condición de los órganos fonatorios) dinámico (relacionado con la actividad de cada uno de ellos) y funcional (se refiere a la producción y uso apropiado de los sonidos) en dos etapas diferentes: la evaluativa y la terapéutica.

El manual en su primera parte: la evaluación, presenta un conjunto de estrategias y reactivos para la evaluación de los problemas de pronunciación. Un análisis de todos los sonidos fonéticos del castellano se constituye en el principal aporte de esta sección. En la segunda parte se identifican los elementos y procedimientos fundamentales para la corrección de estas dificultades. Una programación de actividades secuenciales que se ajustan a las completamente identificadas dificultades del niño y que están apoyadas en los modelos operantes, es el rasgo esencial de esta última unidad.

El amplio conocimiento que sobre el tema posee el autor y el estilo claro y preciso de exposición, hacen que esta obra se constituya en una fuente de consulta útil para todas aquellas personas (padres, maestros, especialistas etc.) que interactúan con un niño dislálico.

Julio Eduardo Cruz

RESEÑADOR DEL LIBRO

JULIO EDUARDO CRUZ. *Afiliación institucional:* Universidad Católica de Colombia. *Grado:* Psicólogo, 1986, Universidad Nacional de Colombia, Bogotá. *Especializaciones:* Psicología social, análisis experimental del comportamiento, psicología educativa. *Dirección de correo:* Carrera 44 No. 67-16, Bogotá, Colombia.

A partir del 1 de octubre de 1988, envíen los manuscritos al próximo Editor:

After October 1, 1988, send manuscripts to the next Editor:

A partir de 1º de outubro de 1988 envie os manuscritos ao seguinte Editor:

José Miguel Salazar
Universidad Central de Venezuela
Apartado 47018
Caracas 1041-A
Venezuela

0400601 • U98M1 • Printed in U.S.A.