ENGINEERING PERSONALIZED INSTRUCTION IN THE CLASSROOM[1]

FRED S. KELLER
Arizona State University
Tempe, Arizona, U.S.A.

This paper outlines a new method of teaching elementary psychology, a method developed by Professor J. G. Sherman and the author at Arizona State University, which aims to provide individualized and personalized instruction in psychology for beginning students, without disrupting seriously the conventional pattern of organization and operation. The method grew out of a series of discussions, in 1962, between Professor Sherman, the author, and two professors at the new University of Brasilia (Professor Rodolfo Azzi and Professor Carolina Martuscelli Bori); and it was first tried out in 1964, at that university, in a one-semester lecture-and-laboratory course.

We began using a modification of the Brasilia procedure at Arizona State in the Spring term of 1965, with some success, as was reported at the Chicago meeting of the American Psychological Association in that year. Since then, we have made further progress along the same lines; and, next year, with the cooperation of Professors Pliskoff and Hegge, we expect to coordinate our efforts, further standardize the technique, and apply it to a larger student population.

The method is being used at present in two sections of an introductory course in which approximately 140 students were initially enrolled. This is a 4-point laboratory course, with class meetings scheduled for 75 minutes on Tuesdays. A 2-hour Saturday-morning session was added early in the course, at student request. The laboratory work of the course is carried out at designated intervals of time and constitutes approximately one-fourth of the course content. There are ten experiments and each is carried out by the student until the principle involved has been successfully demonstrated. This work, which is under the charge of Professor Sherman, will not be discussed further at this time.

The unique features of our method can be presented best, perhaps, by quoting from a mimeographed hand-out which is given to each student on his first day of class-attendance:

"This is a course that permits you to move, from start to finish, at your own pace. You will not be held back by other students or forced to go ahead

[1] This paper was presented, under the title of "Engineering individualized instruction in classroom teaching" at the meetings of the Rocky Mountain Psychological Association in Albuquerque, New Mexico, May 13, 1966.

until you are ready. At best, you may meet all the requirements in less than one semester; at worst, you may not complete the job within that time. How fast you go is up to you.

"The work of this course will be divided into certain *units* of content, which correspond roughly to a series of home-work assignments and laboratory exercises. These units will come in a definite numerical order, and you must show your complete mastery of each unit (by passing a *reading test*) before moving on to the next.

"A good share of your reading may be done in the classroom, at those times when no lectures, demonstrations, or other activities are taking place. Your classroom, that is, will sometimes be a study-hall.

"The lectures and demonstrations in this course will have a different relation to the rest of your work than is usually the rule. They will be provided only when you have demonstrated your readiness to appreciate them; no examination will be based upon them; and you need not attend them if you do not wish. When a certain percentage of the class has reached a certain point in the course, a lecture or a demonstration will be available at a stated time, but your presence there will not be compulsory . . . . .

"You will have an opportunity to ask questions and to discuss the course with your classmates, your *proctor*, your course assistant, or your instructor. Group discussion will depend not only upon your desire, but also upon your preparation—your readiness—to engage in such activity. Questions may be asked in *writing* at any time, and will be individually considered by your instructor.

"The teaching staff of your course will include, as already suggested, a proctor, a laboratory assistant, a classroom assistant, and your instructor. The *proctor* has been chosen for his mastery of the course material and orientation, for his maturity of judgment, for his understanding of the special problems that confront you as a beginner, and for his willingness to assist. He will provide you with all your study materials except your textbooks. He will . . . . . pass upon your readiness tests as satisfactory or unsatisfactory. His judgment will ordinarily be law, but if he is ever in serious doubt he can appeal to the classroom assistant for a ruling, or even to the instructor. Failure to pass a test on the first try, the second, the third, or even later, will not be held against you; better too much testing than not enough, if your final success is to be assured . . .

"Your work in the laboratory will be carried out under the direction of a graduate laboratory assistant, whose detailed duties need not be listed here. In addition, there will be a graduate *classroom* assistant, upon whom your proctor will depend for various course materials (assignments, study questions, special readings, and so on) and who will collect and keep up to date all progress records for all course members. The classroom assistant will con-

fer with the instructor daily, aid the proctors on occasion, and act in a variety of ways to further the smooth operation of the course machinery.

"The *instructor* will have as his principal responsibilities: (a) the selection of all study mtaerials used in this course; (b) the organization and the mode of presentation of these materials; (c) the construction of tests; and (d) the final evaluation of each student's progress . . . It will be his duty also to provide lectures, demonstrations, and discussion opportunities for all students who have earned the privilege; to act as a clearing-house for requests and complaints; and to arbitrate in any case of disagreement or misunderstanding between students and proctors or assistants.

"All students in the course are expected to take a final examination, in which the entire term's work will be represented. With certain . . . exceptions, to be mentioned later, this examination will come at the same time for all students, at the end of the term. [The exceptions were those students who finished early and chose to be examined immediately, and those students who took a grade of Incomplete.] The examination will consist of questions which, in large part, you have already answered in your readiness tests. Twenty-five percent of your course grade will be based upon this examination; the remaining seventy-five percent of your grade will be based upon the number of units of reading and laboratory work that you have completed successfully during the term."

Along with this description and a few instructions, the student was given his first assignment and his course began. The main classroom was used as a study hall; a smaller classroom was used for test-taking; and a third room, containing seven proctors' cubicles, was set aside for grading and discussion of tests. When a student in the study hall felt ready for examination, he went to the test room, where the assistant gave him a set form and his exam book. When the questions were answered, the student returned the test form and took his exam book to his proctor. After the test was evaluated, it was retained by the proctor, and the student either received a new assignment or was sent back to study for another test on the same material, usually after considerable discussion of the errors made. The graded test was passed on to the classroom assistant or the instructor at the end of the hour, to be checked through and recorded.

As part of the reading assignment for each course unit, the student is given a list of "study questions," perhaps thirty to forty in number, which cover every element of the reading for which he will be held responsible. If he can find the answers to these questions, and remember them, he will come to the testing room with a high probability of passing the first quiz that he takes on any unit.

The readiness tests themselves are commonly composed of ten to twenty questions of the completion type, along with one short-answer essay question.

There are four or five alternate test forms for each unit, but the average number required for passing a test with a perfect score is probably no more than two. After the fifth course-unit has been mastered, each new-unit test is accompanied by a review test. Thus, the Unit-6 readiness test is accompanied by a review for Unit 2; and so on.

On every Monday afternoon during the term, the proctors meet with the assistant and the instructor, to go over the preceding week's work and prepare for the next. At this time, new test questions and their answers are discussed in detail, and a considerable amount of incidental teaching takes place. We have worked out informal rules for grading, so that our procedures will have maximal consistency from proctor to proctor; and we are currently conducting an item analysis of all tests, with the aim of eliminating bad items, which seem to turn up regularly in spite of years of experience in test-making. The number of different interpretations that may be given to a perfectly clear and unambiguous question continues to be appalling. "Good errors" are as real for college students as they were for Professor Kohler's chimpanzees. When "perfection" is required of each student before permitting him to move on, one comes to look upon quizzing in a brand new light.

The selection of proctors is made from undergraduates who have taken the same course, or a course having the same content, in an earlier term. Originally, we wanted junior or senior majors in psychology, with graduate-school intentions, but we do not have enough of these to satisfy our present proctor-student ratio of one to ten. We have been forced to take non-majors, sophomores, and, in one instance, a second-term freshman. A well-motivated A-student, without obvious personality defects, and under adequate supervision, is likely to meet the requirements of the job, without serious trouble.

The first proctors in the course were volunteers from the top 20 of a conventional lecture course having the same general content as our present one. Initially they worked without pay, but were later given a nominal hourly wage. This has now been discontinued, and the present group receives two points of course credit—which they more than earn.

Lectures, demonstrations, and an occasional movie are announced in advance, along with the unit requirement for attendance, and every effort is made to have them as rewarding as possible. By common standards, they are short and infrequent. Even so, they will not compete successfully against test-taking except in the case of a few students, and it was necessary to stop all testing while such events are going on.

A chart of student progress through the reading units of the course is maintained at all times and is on regular display. Every student knows where he is, and where he ought to be, with respect to work accomplished—at least he knows if he attends the class. The burden of responsibility for study and test-

taking is placed solely upon the student himself. No goading, no exhorting, no chiding, no promising or threatening of any sort is provided at any time. If a student wants to pass the course, he knows what to do; and he will be rewarded therefor, but he will not be babied.

At the Chicago meetings, it was pointed out that our results thus far with this method have been consistently encouraging. About 90% of our students say that, although the method requires more work than do the usual course procedures, they *learn more*, they feel a *greater sense of achievement*, they find *greater recognition as individuals*, and they *enjoy the work more* than in conventional courses. They also report, quite generally, an *improvement in their study habits* as the course goes on, along with *a more positive attitude toward test-taking*. Many of them have expressed the wish that some of their other courses could be taught in the same manner.

The proctors and assistants have also been enthusiastic about the method and its educational value, for themselves as well as their charges. They claim to have profited greatly from preparing for, and giving, readiness tests, talking with students about the assignments, and discussing the work to come with their assistant or instructor. Confirming such claims is the fact that the proctors are almost never absent from their classes or their Monday briefings; that they regularly volunteer for Saturday-morning duty; and that they would gladly be proctors again, for another tour of duty.

The basic method might, of course, function without them. This was the case during the first term in Brasilia, where a paid employee did all of the testing; and Professor Sherman, with an assistant, did all his own proctoring during his first-semester try-out of the method. One can even imagine that the course could be computerized, or carried out in the correspondence-school pattern. But I am at this time convinced that, with proper supervision and moderately careful screening, *for the task at hand*, these proctors constitute one of the best features of our course—if not *the* best. They are the ones who provide, for each individual student, that personal contact which is commonly denied to all but a select few in most of our classes, and which is especially meaningful for the beginner in any field. For this beginner, the instructor, and even the assistant, may be a far-off figure in another world, but the proctor is a person of status, who has the answers, understands the problems, and is willing to help. The relation between the student and the proctor is established more firmly and workably with every test that is graded—whether it be a *pass* or *fail*. And every step forward made by the student is almost as satisfying to the proctor as it is to the student. A better system of mutual reinforcement is hard to picture.

The *instructor* in such a course may work harder, at least in *establishing* the course, than he has done before for the same points of academic load, but

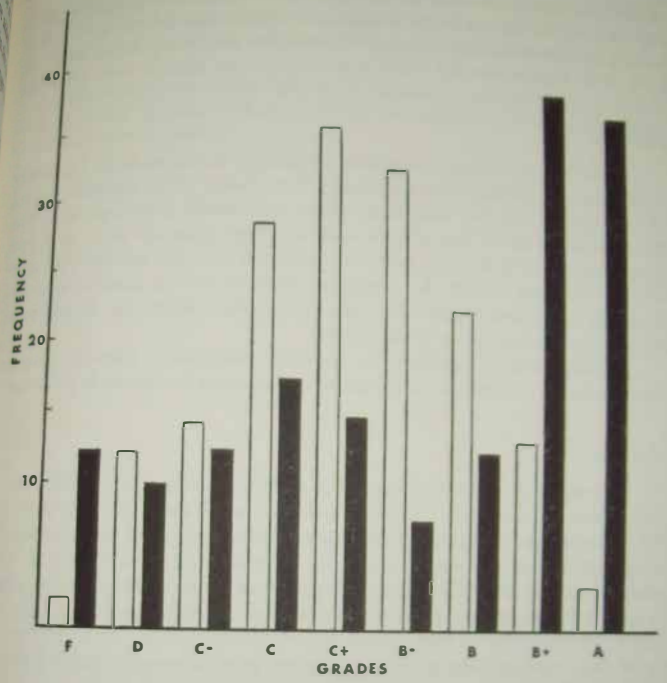
he profits more. He finds a new appreciation of his labors and a new dignity in his profession. He is no longer dogged with the feeling that he gives his all to those who *don't understand it, don't want it,* and *don't deserve it.* Someone has said that learning is *not a spectator sport.* He might have added, with equal truth, that *many teachers treat it as if it were.* The teacher in *this* course, however, in his short, infrequent classroom appearances no longer feels like a second-rate entertainer or that he is talking to himself.

Other aspects of the teacher's job become more important than his public appearances. Initially, he must state his goals, in fairly explicit form. He must make a tentative analysis of his course content, and put it into suitable units of instruction. He must know his chosen textbook material more intimately perhaps than he has ever known it before. He must construct an exhaustive variety of study questions and an *exhausting* number of test questions. These test questions must be good ones—as unambiguous as possible, as fair as he can make them, not tricky, humorous, or show-offish. If he designs them well enough, they will help him teach. He must learn to write clear supplements to unclear or over-difficult reading assignments, thus improving his own expository style and perhaps gradually building a text of his own. He must be alert to the progress of every student—the slow as well as the fast—and he must keep a sharp eye on the work of his subordinates. The instructor is a very busy man, but it is a very satisfying busy-ness.

During the present semester, a copy was secured of a mid-term examination given during the first half of a one-year introductory course at Columbia University, in 1964, along with the grade distribution for the 191 students who took this exam. It was a 100-point, objective-type examination and, after subtracting twelve or thirteen points' worth of questions relating to material not covered in our own course, the remainder was given as an unnounced quiz to the 66 students who were in attendance on the mid-term date at A.S.U.

After correcting for the difference in number of questions and number of students, I obtained the results represented in the figure. The empty bars are for the Columbia distribution, the filled ones for A.S.U. The highest, as well as the lowest, of the grades are for A.S.U. students. Fifteen of them, out of the 66, received A's, eight of which were higher than the Columbia A's. One A.S.U. student scored 100. Five A.S.U. students of the 66 failed, three with scores much lower than the Columbia F's. (Since our quizzes were unsigned, it cannot be determined if these F's are related to the distance traveled in the course at the time of the test.)

The A.S.U. distribution is unusual, but typical of all the distributions collected thus far. It would be more extreme if an examination were given today, with more A's and B's, perhaps more F's, and fewer C's and D's. The distribution could, of course, be changed by decreasing or increasing the difficulty of



the unit assignments, but only if students were denied the privilege of completing the course in a later semester.

There are obvious faults with this comparison. More of our poor students were probably absent on the day we gave the test; our quiz was unannounced, which may have favored Columbia; the questions were Columbia-prepared, etc., etc., but the results are at least suggestive.

The work done thus far to develop this kind of individualized teaching within the mass-education framework of a large university has been aimed mainly at establishing a base-line procedure. Once this is done, it will be possible to begin assessing the relative importance of some of the variables involved, and introducing others into the teaching situation. Our procedures

are still imperfect; our reach is uncertain; and no claim to originality is made. But the method does achieve a degree of educational control greater than before, without sacrificing a personal interaction which may, in the long run, have more significance for the educational process than is now perceived.

ABSTRACT

A method of teaching elementary psychology developed at the University of Brasilia and Arizona State University. The content of the course is divided into ten units, which the student completes individually at his own pace. The student is tested at the end of each unit, and must score 100 per cent correct before passing on to the next unit. The results of each test are discussed individually with the student by a proctor, who is another student who has previously taken the course. This one-to-one relationship may be the most important part in the success of the method.

A comparison of results on a mid-term examination between Arizona State University students and students from Columbia University is presented.

RESUMEN

Un método que ha sido desarrollado en la Universidad de Brasilia y en la Universidad del Estado de Arizona. El contenido del curso se ha dividido en diez unidades que el alumno completa en forma individual sin apremio de tiempo. Al final de cada unidad se efectúa una evaluación de su trabajo. En esta prueba el alumno debe obtener un cien por ciento de respuestas correctas antes de pasar a la otra unidad. El alumno discute los resultados de cada prueba en forma individual con un tutor que es otro alumno que tomó el curso anteriormente. Esta relación frente a frente puede que sea el elemento decisivo hacia el éxito del método.

Se presenta una comparación de los resultados de una prueba parcial entre alumnos de la Universidad de Arizona y los de la Universidad de Columbia, en Nueva York.

RESUMO

Um método para o ensino de psicologia elementar desenvolvido na Universidade de Brasilia e na Universidade do Estado de Arizona. O conteúdo do curso é dividido em dez unidades, as quais são completadas pelo aluno a seu próprio passo. Ao fim de cada unidade o aluno faz um exame sôbre o material estudado, e, se obtem 100% de acertos, recebe a unidade seguinte. Os resul-

tados de cada prova são discutidos individualmente pelo aluno e um orientador, sendo êste último um outro aluno que já fêz o curso. Esta relação pessoal entre orientador e aluno, talvez seja o fator mais importante, relativamente ao sucesso do método.

Aprensenta-se uma comparação entre os resultados da prova de metade do curso, de alunos da Universidade do Estado de Arizona e da Universidade de Columbia, respectivamente.

EL NIVEL DE CONCEPTUALIZACION: UNA CLASE DE RESPUESTAS VERBALES CONTROLADAS POR LAS CONTIGENCIAS DE REFORZAMIENTO

LÓPEZ R., FLORENTE Y RIBES I., EMILIO.
Universidad Veracruzana. Xalapa México

INTRODUCCIÓN

Tradicionalmente, la conducta conceptual ha sido manejada en un nivel casi puramente descriptivo, y la investigación experimental de los procesos que subyacen a la formación de conceptos ha sido inferida a partir de estas descripciones. Ejemplo de esto último, puede ser la clasificación de los conceptos en conjuntivos, disyuntivos y relacionales, y los tipos de estrategias de selección y recepción que son inferidos a partir de las elecciones del sujeto (4), o los estudios clásicos en términos de nivel conceptual (7, 10). Sin embargo, estos enfoques han resultado poco fructíferos, ya que el primero carece de solidez teórica y los segundos hasta ahora han resultado meramente clasificatorios. Más aún, la tendencia actual es entender la conceptualización, no como una sola dimensión conductual, sino como varias dimensiones reguladas por distintos principios funcionales (1, 5, 8, 12).

Recientemente, el estudio de los conceptos ha recibido un fuerte impulso al intentar acercarlo a los principios estímulo-respuesta. Berlyne (2), en una interpretación de la equivalencia de estímulos, propone que ella está basada en un proceso en el que subyacen la generalización primaria y secundaria de estímulos, siendo característica de la conducta conceptual la generalización secundaria del estímulo y los sistemas de respuestas nominativas (11). Por otra parte, la importancia de los mediadores verbales ha sido señalada por Kendler (8) en la solución de problemas y conceptualización, haciendo notar que la interpretación de sus resultados en un estadio resultaría inadecuada.

En estos términos, López R. (9) realizó una investigación social, tomando en cuenta las dimensiones utilizables para conceptualizar un objeto y las respuestas verbales que con mayor probabilidad serían condicionadas por una clase social particular (alta o baja). Se predijo, en base a las respuestas concretas y las funcionales, que la clase alta daría preferentemente respuestas del primer tipo y la clase baja del segundo. Los resultados fueron confirmados en las respuestas dadas a una prueba de semejanzas, pero no resultaron significativos los datos obtenidos simultáneamente en la prueba de Clasificación de Objetos de Gardner (6).

Hipótesis.—Lo anteriormente descrito, nos hizo pensar en la posibilidad de condicionar experimentalmente respuestas de tipo funcional o concreto ante grupos de objetos. En otros términos, consideramos que el nivel preferido de conceptualización podría ser condicionado bajo circunstancias experimentales, ya que el hecho de que existan preferencias de nivel de abstracción en las clases sociales es indicativo de entrenamientos diferenciales determinantes de la probabilidad de ciertas respuestas ante grupos de estímulos específicos. En esta forma, los procesos pueden ser aclarados si dichas respuestas pueden darse ante una situación de laboratorio bajo entrenamiento experimental, en donde se controlen las contingencias de reforzamiento.

Método.—Se eligieron veintiocho Ss al azar, todos ellos alumnos de la Escuela Primaria anexa a la Escuela Normal Veracruzana de Xalapa, siendo catorce niños y catorce niñas, cuyas edades fluctuaban entre los ocho y los doce años de edad. Se les administró una variedad verbal de la prueba de Clasificación de Objetos de Gardner (6), elaborada por los autores, consistente en una hoja con los dibujos de los siguientes objetos de uso común: un martillo, un jabón, un puro, una paleta de albañil, una pipa, un peine, una escobetilla, una corbata de moño, una cuchara, un gotero, un cepillo, un lápiz labial, un cigarro, una horquilla para el pelo, unos anteojos, un cuchillo, un tenedor, una cápsula, un arete, un frasco de aspirinas y un serrucho.

Las instrucciones para resolver la prueba eran: "Primero que nada, quiero que sepan que esta prueba no tiene contestaciones buenas o contestaciones malas. Cada quién la resuelve a su manera. Lo que ustedes tienen que hacer es poner un mismo número en la raya que hay abajo de cada dibujo, en todos aquellos objetos que ustedes crean que deben ir juntos. Pueden poner todos los objetos que quieran juntos, siempre y cuando exista una razón para hacerlo. Aquellos dibujos que no vayan juntos con ningún otro pueden ponerles abajo un número diferente. Después de esto, en la hoja de atrás, en el renglón que corresponde a cada número que ustedes hayan puesto en los dibujos, deben poner por qué razón los pusieron juntos." En una hoja anexa a la de la prueba, los Ss tenían que escribir el motivo por el que habían agrupado dichos objetos. Los registros obtenidos en la primera aplicación de esta prueba, fueron utilizados como línea base para poder evaluar posteriormente los cambios tentativos bajo la administración de reforzamiento. El siguiente paso consistió en dividir a los Ss en dos grupos homogéneos al azar y proyectarles transparencias de los objetos dibujados en la prueba verbal, en las que eran representados con un ser humano en un contexto habitual: un hombre con un martillo, una mujer con un arete, etc. A uno de los grupos se les reforzó el emitir respuestas verbales de tipo funcional (implicando utilización) en relación a los objetos, mientras que el otro grupo lo fué con respecto a respuestas verbales de tipo concreto (carac-

terísticas sensoriales del objeto). El reforzamiento fué programado intermitentemente de acuerdo con un programa de razón variable. El reforzamiento consistía en una combinación de dinero, que se administraba a los niños intermitentemente cuando ellos repetían propiedades funcionales o sensoriales de los objetos proyectados en la pantalla, que uno de los experimentadores modelaba en voz alta para tales propósitos. Inmediatamente después de terminar esta operación, se les aplicó nuevamente la modificación verbal de la prueba de Clasificación de Objetos de Gardner, a fín de observar cambios en la línea base de respuestas conceptuales. Se computaron además de las respuestas funcionales y concretas, el número de grupos formados y las respuestas designadas como "otras" (respuestas abstractas, objetos aislados, confabulaciones).

Análisis de los Datos.—Se aplicó la prueba t para averiguar el nivel de confiabilidad de las diferencias entre las medias de las frecuencias de respuesta obtenidas del primero y segundo registros con la prueba verbal de Clasificación de Objetos (ver Tablas I y II).

No encontramos ninguna diferencia significativa, $p > 0.10$, entre todas las comparaciones posibles de medias de frecuencia de respuestas o número de

TABLA I

Medias de frecuencia de respuesta del grupo al que se reforzó nivel funcional, antes y después del tratamiento experimental, N=14.

Clase de Respuesta	Test X	ReTest X
1. Núm. de Grupos	8.26	9.31
2. Resp. Funcionales	2.42	2.47
3. Resp. Concretas	2.63	1.53
4. Otras	2.47	4.05

TABLA II

Medias de frecuencia de respuesta del grupo al que se reforzó nivel concreto, antes y después del tratamiento experimental, N=14.

Clase de Respuesta	Test X	ReTest X
1. Núm de Grupos	8.47	8.35
2. Resp. Funcionales	2.52	2.35
3. Resp. Concretas	1.47	1.05
4. Otras	4.41	4.47

grupos. Se efectuó primero la comparación de los registros del grupo al que se le reforzó la emisión de respuestas funcionales así como al que se reforzaron las respuestas referentes a propiedades sensoriales del estímulo. Después se analizaron las medias por edad, por sexo, por número de grupos, por respuestas computadas como "otras." No fueron significativas ninguna de las diferencias obtenidas entre el registro de respuestas después de la administración del reforzamiento con respecto a la línea base obtenida antes del tratamiento experimental.

Discusión.—Los resultados obtenidos no apoyan la hipótesis expresada en un principio, aún cuando algunos de los datos varían en la dirección predicha. Sin embargo, no consideramos que ésto se deba básicamente a la variable independiente manipulada experimentalmente o que pueda interpretarse como una confirmación gruesa de la hipótesis inicial. Existen varias razones por las que creemos que no se produjeron los resultados esperados. En primer lugar, el registro de los datos sobre nivel conceptual fué colectivo, lo que hizo imposible un control estricto en la forma de resolver la prueba verbal de Clasificación de Objetos. Es decir, no existe evidencia definitiva en el sentido de que las instrucciones hayan sido cumplidas o siquiera comprendidas. En otras palabras, ni la línea base de respuesta, ni los cambios en la misma obtenidos en la segunda sesión, son confiables en nuestra opinión por lo anteriormente mencionado. En segundo lugar, Bijou (3) ha encontrado que las respuestas conceptuales deben reforzarse todas y cada una de ellas, utilizando un programa de reforzamiento contínuo, pues los programas de reforzamiento intermitente o parcial no producen los efectos deseados. Tal como ya lo hemos mencionado, el programa utilizado fué de razón variable y ello puede explicar en gran medida los datos obtenidos. Finalmente además de existir la posibilidad de que no se hayan efectuado los ensayos necesarios para establecer sólidamente la respuesta (solamente se reforzó una vez el nuevo nivel de respuestas), las contigencias de reforzamiento al administrarse éste en grupo, desconocemos si fueron o nó relevantes a la respuesta verbal o a alguna otra de las conductas manifestadas en la situación por los Ss (estar en el salón, hablar, etc.). La carencia de otros controles sobre la conducta emitida por los niños en la condición experimental, nos impide verificar esta posibilidad. Por lo tanto, aún cuando los resultados no siguieron el curso predicho por la hipótesis, se obtuvieron datos valiosos que utilizaremos para modificar el procedimiento experimental en investigaciones ulteriores. Interpretamos el estudio, no como evidencia en contra de la aplicación de los principios estímulo-respuesta a la conducta conceptual, sino más bien como una experiencia piloto que nos ha de permitir en el futuro un control más riguroso de las variables involucradas en la conducta conceptual. Estudios subsecuentes explorarán la misma hipótesis bajo condiciones de prueba individual, en las cuales esperamos poder verificarla plenamente.

REFERENCIAS

(1) Archer, E James. The Psychological Nature of Concepts en Klausmeier, H. y Harris, Ch. W. (Eds.) *Analyses of Concept Learning*. New York: Academic Press, 1966, pp. 37–50

(2) Berlyne, D. E. *Structure and Direction of Thinking*. New York: John Wiley, 1965.

(3) Bijou, Sidney W. Comunicación Personal. Febrero, 1967.

(4) Bruner, J. S., Goodnow, J. J. and Austin, G. A. *A Study of Thinking*. New York: John Wiley, 1956.

(5) Gardner, R. W., Jackson, D. N. and Messik, S. J. *Personality Organization in Cognitive Control and Intellectual Abilities*. New York: International Universities Press. Psychological Issues, 1960, Vol. II, No. 4.

(6) Gardner, R. W. and Schoen, R. A. Differentiation and Abstraction in Concept Formation. *Psychological Monographs*, 1962, vol. 76, No. 41.

(7) Goldstein, K. and Scheerer, M. Abstract and Concrete Behavior: An experimental Study with Special Tests. *Psychological Monographs*, 1941, vol. 53, No. 2.

(8) Kendler, H. H. and Kendler, T. S. Vertical and Horizontal Processes in Problem Solving. *Psychological Review*, 1962, 69, pp. 1–16.

(9) López R., Florente. Algunos Factores en el Condicionamiento Social de Respuestas de Nivel Conceptual. *Memorias del Primer Congreso Mexicano de Psicología*. Xalapa, Ver., marzo, 1967.

(10) McGaughran, L. S., and Moran, L. J. "Conceptual level" vs. "conceptual area" analysis of object-sorting of schizophrenic and non-psychiatric groups. *J. abnorm. soc. Psychol.*, 1956, 52, 43–50.

(11) Osgood, Ch. E. *Method and Theory in Experimental Psychology*. New York: Oxford University Press, 1953.

(12) Staats, A. W., and Staats, C. K. *Complex Human Behavior*. New York: Holt, Rinehart & Winston, 1963.

RESUMEN

Se formuló una hipótesis sobre la posibilidad de aplicar los principios estímulo-respuesta al estudio de la conducta conceptual. Se administró un programa de reforzamiento de razón variable para averiguar si podrían cambiarse los niveles de respuesta conceptual funcional y concreto de un grupo de niños en edad escolar, tal como se miden dichas respuestas por una versión verbal de la prueba de Clasificación de Objetos de Gardner. Los datos no confirmaron la hipótesis. Sin embargo, se proponen argumentos que permitan explicar los resultados sin desechar la hipótesis inicial, proyectándose las condiciones futuras de investigación del problema.

ABSTRACT

An hypothesis was formulated concerning the possibility of applying S-R

principles to the study of conceptual behavior. A program of variable ratio reinforcement was employed with a group of school children to determine if the levels of conceptual response, functional and concrete, could be changed, as such responses are measured by a verbal version of the Gardner Object Classification Test. The data did not confirm the hypothesis. Nevertheless, arguments are presented to explain the results without rejecting the initial hypothesis, which suggests possibilities for future research on this problem.

RESUMO

Formulou-se uma hipótese sôbre a possibilidade da aplicação do princípio de estímulo-resposta ao estudo do comportamento conceitual. Usando-se como sujeitos crianças de idade escolar, administrou-se um programa de reforçamento de razão variável para se verificar a possibilidade de mudança dos níveis de respostas conceituais funcionais e concretas. Tais respostas forma medidas através de uma versão verbal do teste de classificação de objetos de Gardner. A hipótese não foi confirmada. Os resultados obtidos são discutidos com vistas a futuras abordagens experimentais do problema.

RIGOR VERSUS VIGOR: SOME DUBIOUS ISSUES IN THE DEBATE ON RESEARCH PHILOSOPHY[1]

HERBERT C. KELMAN
The University of Michigan, U.S.A.

Social and behavioral scientists devote a considerable amount of their time and energy to discussions about alternative philosophies of research. Such discussions are especially prominent in interdisciplinary research settings, where they often take the form of debates between the disciplines. For example, during the past fifteen years, research in mental hospitals has drawn on the contributions of anthropologists, psychiatrists, psychologists, and sociologists. Investigators coming out of these different research traditions have tended to formulate the research problems in different ways, to focus on different units of analysis, and to favor different research methods. These differences tend, on occasion, to take on ideological overtones and to degenerate into contests over the relative virtues of the different disciplines and over their relative faithfulness to the precepts of science or to the phenomena under study.

Debates about alternative philosophies of research may also take place *within* a given discipline. Thus, for example, within the field of political science, a debate between "behavioralists" and "traditionalists" has been going on at least since the end of World War II. The behavioralists favor quantitative research and draw their concepts from such fields as psychology and sociology. The traditionalists are more inclined to follow historical, descriptive, and normative approaches to their subject matter. Again the debate has often tended to be ideological, with each side questioning the very legitimacy of the enterprise in which the other side is engaged.

Within the field of social psychology, which is interdisciplinary by its nature—being rooted in psychology and sociology, and influenced by other disciplines as well—it is inevitable that differences in research philosophy would arise. These differences are in part related to the disciplinary origins—i.e., psychology or sociology—of those who call themselves social psychologists. Many of the differences, however, cut across these disciplinary lines and arise as much within the two subsets of social psychologists as they do between them.

Social psychologists differ in their view of the proper *level of analysis* for

[1] This paper will appear as a chapter in a book on The Value Context of Social Research, to be published by Jossey-Bass, Inc. (San Francisco) in the fall of 1967. It is based, in part, on a paper that was originally prepared while the author was working at the National Institute of Mental Health (U.S.A.).

social-psychological research. Thus, some investigators (including some who were originally trained in sociology) regard social psychology as essentially an extension of general psychology. They attempt to analyze social behavior in terms of general psychological principles, operating under the special conditions provided by the social situation. Their strategy is to take as their starting-point models developed for the analysis of simple situations and to see how far these can be pushed in the analysis of the more complex situations in which social behavior occurs. Other investigators (including many who were originally trained in psychology) assume that social behavior cannot be effectively reduced to the level of individual behavior and that social psychology must, therefore, develop its own level of analysis. Their strategy, accordingly, is to start out with models that represent—as parsimoniously as possible—the complexity of the phenomena that they hope to illuminate.

Social psychologists also differ in the *research methods* that they prefer to use. Roughly speaking, one might distinguish between those whose orientation is primarily quantitative and those whose orientation is more equalitative—or at least clinical and non-statistical.[2] Within each of these two sets of orientations, one can make further distinctions, which often reflect important differences in research philosophy. Thus, among the "statistically" oriented social psychologists, there are those who rely primarily on the experimental method, involving the active manipulation of variables to be studied and observation of their effects; and those who rely primarily on the use of survey methods to obtain opinion data and behavior reports from selected samples of the population. Similarly, among the more "clinically" oriented social psychologists, one can distinguish between those who prefer to use structured methods to obtain their data on individuals and groups, such as personality tests, structured interviews, and sytematic group observation; and those who prefer to use the methods of participant observation and community study in the anthropological tradition.

As in the case of other differences in research philosophy that I have touched upon, the differences among social psychologists are sometimes ideologized. Proponents of one orientation may insist that they alone are truly scientific; proponents of another orientation may counter that they alone are truly relevant. Debates about research philosophy are a healthy and necessary part of any field of investigation, but when they take the form of arguments about the *general* virtue of "my" concepts and methods as compared to "your" concepts

[2] These terms are only approximations. I would include in the second category studies in the factor-analytic tradition that collect large amounts of data on a single individual or a small number of individuals. These studies are certainly quantitative, in the sense that they work with numerical scores and subject these to various kinds of mathematical operations. They are, however, non-statistical, in the sense that they are not primarily concerned with drawing inferences on the basis of sampling procedures (whether these be random sampling from a surveyed population or random assignment of subjects to experimental and control groups).

and methods—without regard to the specific uses to which these are put—then they are bound to be fruitless. More often than not, when we argue about the relative merits of psychological vs. sociological, or quantitative vs. qualitative, or experimental vs. naturalistic approaches, we are dealing with false issues.

Sometimes the issues are false because they are based on a misconception of what the "opposing" approach or its particular representative stands for. Thus, antagonists in the debate between psychological vs. sociological orientations may be operating with stereotyped and anachronistic notions of how the other discipline approaches its problems. I have heard psychological analysis criticized because, supposedly, it seeks to explain group or collective behavior in terms of the idiosyncratic personality characteristics of individuals. I have heard sociological analysis criticized because, supposedly, it is based on armchair speculation or at best on casual observation. Such conceptions may, perhaps, have been accurate at one time, and they may still be good characterizations of how some psychologists or some sociologists go about their repective tasks. But they are highly distorted pictures of the dominant trends in these disciplines, and one can certainly not assume that they are accurate characterizations of the approach of any *given* psychologist or sociologist. Above all, such stereotypes ignore the fact that within each of the two disciplines, there is wide diversity in the concepts, methods, and levels of analysis used.

Similar misconceptions may be at the root of certain other dubious controversies. For example, some may be critical of a qualitative approach because they see it as an undisciplined selection of illustrations that can support any point the investigator wishes to prove. While we have all seen research that conforms to these unhappy stereotypes, they are clearly caricatures and perversions of the skilled and intelligent use of quantitative or qualitative analysis. Whenever debaters operate on the basis of such mutual misconceptions, there is no possibility for a true confrontation and resolution of issues. The antagonists have not even come to the point of discovering whether a real difference exists between them.

Even when two antagonists are genuinely talking to each other and real differences in research approach do exist between them, they may be debating false dissues. This happens when the debate contrasts two alternative approaches that are not really alternatives at all. The debate assumes that they are two different ways of doing the same thing, when in fact they turn out to be two different ways of doing two different things. It makes no sense to argue whether one set of concepts or methods is better than another when the two are designed to deal with different kinds of problems. The question must always be: better for what? Two investigators—both of whom may, for example, be studying social interaction—may be completely justified in working with different variables and units of analysis, if one is primarily concerned with the

effects of interaction on individual personality and the other with its effects on the social system. Similarly, the investigator concerned with defining the contours of a phenomenon cannot reasonably be criticized for his failure to use an experimental approach, just as the investigator concerned with establishing functional relationships between variables cannot be criticized for his failure to use a more naturalistic approach. Debates that ignore the purposes that the competing approaches are designed to serve cannot provide much illumination.

The real issue is whether a particular research approach is appropriate to the questions that the investigator is trying to answer. Beyond that, there is the more complicated matter of whether these quesions are worth asking at all. This, of course, raises the larger issues of the definition of the social researcher's task and the nature of the enterprise in which he is engaged. These real issues —concerned with the fit between our research approaches and our research problems, and with the nature of our essential task—form the background for much of the debate on research philosophy that is taking place. They tend to be masked, however, and relegated to the periphery, when the discussion focuses on false dichotomies, rooted in competing disciplinary loyalties and scientific models.

Many of the dubious issues that I have already mentioned turn on one or another aspect of what might be called the "rigor vs. vigor" controversy. In the pages that follow, I shall try to spell out why the controversy on this level is of questionable validity, and how some of the real issues in the evaluation of social research cut across this particular dichotomy. I shall then turn to one special type of research approach, which is generally seen as the model of rigor—namely the experimental method. By examining experimental work in the context of what I regard as the essential task of the social scientist, I hope to demonstrate that, even in this case, the rigor vs. vigor formulation provides a misleading picture.

The Rigor Versus Vigor Controversy

Some years ago, Kenneth Boulding took certain liberties with the lyrics of W. S. Gilbert to drawn portraits of the scientist and the humanist. With a few further modifications, Boulding's lyrics can serve to demonstrate some of the differences between the proponents of rigor and the proponents of vigor. With apologies to both Gilbert and Boulding, then, let me present this brief debate between representatives of the two schools of thought:

R. Conceive me if you can
A rigorous young man,
A physical-causative,
Logical-positive,

White-coat-and-rat young man,
Who has decided bent
Towards experiment,
And Oh! what a wrath is his
If a hypothesis
Claims it's self-evident.

V. Conceive me if you can
A vigorous young man,
A most analytical,
If not always critical,
Raw-stuff-of-life young man,
Who is persuaded that
Man differs from the rat.
To pierce men's reality
He views their totality
In their native habitat.

R. A serious-aims young man,
A Theory-of-Games young man,
A very stochastical
Iconoclastical,
Testable-claims young man.

V. A very unique young man,
A reach-for-the-peak young man,
A flagrantly mystical
Most unstatistical
Rather oblique young man.

R. A one-way-screen young man,
A keep-it-clean young man,
A fine-observational,
Most operational,
Truth-is-what's-seen young man.

V. A truly all-round young man,
Yet very profound young man,
An I'll-only-do-it-if-
It-is-intuitive
Truth-at-a-bound young man.

R. A look-at-the-fact young man,
A get-out-and-act young man,
A set-up-a-project-

Where-money's-no-object,
Foundational-backed young man.

V. A very arm-chair young man,
A no-questionnaire young man,
A find-it-out-many-ways,
Know-it-all-anyways,
Devil-may-care young man.

R. A facts-by-the-yard young man,
A punchable-card young man,
Statistical-tabular,
Special-vocabular,
Work-very-hard young man.

V. A fitful-gleam young man,
A beautiful-dream young man,
A most metaphorical
Highly rhetorical
Letting-off-steam young man.

Needless to say, these portraits are stereotyped caricatures of the two sides, but they do transmit some of the flavor of the debate between them. The proponents of rigor stress the importance of obtaining hard facts through the use of exact methods. They prefer research situations that are maximally structured and, if at all possible, based on the experimental manipulation of the independent variables; research instruments that are impersonal and do not require too much filtering through the mind of the investigator; and dependent variables that can be stated in terms of quantitative indices and subjected to statistical analysis. Objectivity, precision, and replicability are, in their view, the central requirements of a scientific analysis. The proponents of vigor, on the other hand, stress the importance of research methods that will capture the real-life flavor of the phenomenon under study, in all of its fullness and richness. They prefer to make their observations in natural settings, free from the manipulations of the investigators; to use themselves as active instruments for sorting, understanding, and integrating what they observe; and to present their findings in ways that properly reflect the totality of man and society and the complexity of social behavior. Precision of methods is, in their view, less crucial than relevance to the questions that social science proposes to answer and truthfulness to the phenomena of social life.

The different components of the two syndromes I have described do not necessarily always go together, nor do most social psychologists clearly identify themselves with one or the other of these syndromes. Yet this division is at

least an undertone in much of the debate and position-taking on questions of research philosophy and of the most productive direction for the future development of the field.

The question of whether experimental or naturalistic approaches are ultimately superior is meaningless, because it cannot possibly be answered in general terms. The choice of appropriate methods depends, first of all, on the nature of the problem under investigation—the kinds of questions that the investigator is asking. A fetishistic insistence on rigorous methods is bound to close off many promising sources of insight and information. Conversely, a squeamish reluctance to pin things down lest the sanctity of the phenomenon be destroyed is likely to keep us from ever attaining reliable knowledge. Depending on our purposes, a sacrifice in either precision or naturalness may be completely justified. For example, if we are interested in identifying the variables that define a particular phenomenon and in tracing its course of development, then we have to observe it in its natural setting. A less rigorous and more impressionistic approach is clearly called for here, since it yields the data we need and cannot obtain in any other way. On the other hand, if we are interested in establishing functional relationships between two sets of variables, then there is no substitute for an experimental approach, even though this forces us to rely on artificial situations of limited generalizability.

The choice between experimental and naturalistic research depends also on the stage of development of our research area and of the specific problem on which we are working. With respect to the research area in general, the "scientific ethos" requires us to use the best methods available for dealing with the questions we have posed. The unavailability of precise methods is certainly no reason for turning away from these questions. We tackle them as best we can, noting the limitations of our methods and gradually working toward their improvement.

With respect to the stage of development of the specific problem, different approaches are called for at different points in time. In the initial stages of work on the problem, when the investigator is interested in gaining an intuitive understanding of the phenomenon and in developing hypotheses, a clinical-impressionistic approach is most appropriate. When he is ready to test these hypotheses in the form of functional relationships, an experimental approach or some approximation thereof is usually called for. In a later stage, when the investigator is interested in testing the generality of his hypotheses in real-life settings, he may again turn to naturalistic observations or to the use of survey methods.

Finally, the choice of experimental or naturalistic approach must also depend on the preferred style of the investigator. Some are more comfortable with one approach and some with another—for reasons of personality, train-

ing, or esthetic preference. When an investigator uses the approach that is most congenial to him, he is likely—other things being equal, of course—to be more creative and to make a better contribution. Thus, what is a productive method for one investigator is not necessarily so for another. Each investigator should feel completely free to select his own preferred style—without, however, claiming that it is therefore objectively better for all problems. He should keep in mind the limitations of the particular approach he has selected, apply it only to those problems to which it is naturally suited, and derive from it only those conclusions that it is capable of producing. In short, if he writes musical comedy, he should not try to pass it off as epic drama.

Two aspects of the rigor vs. vigor issue deserve special comment. One is the question of quantitative vs. qualitative analysis; the other the question of a holistic vs. an elementalistic approach.

Qualitative analysis is sometimes criticized on the grounds that the investigator has no conception of what he has really found and is highly susceptible to the fallacy of the positive instance. Quantitative analysis, on the other hand, may be criticized for the investigator's tendency to let the true phenomenon pass him by while he is simple-mindedly and obsessively counting the irrelevant. Thus, critics of a qualitative approach argue that the conclusions derived from it may be interesting, but are probably not true; while critics of a quantitative approach argue that the conclusions derived from the latter may be true, but are probably not interesting.

The mere posing of the issue in terms of qualitative vs. quantitative is bound to be misleading. In a very basic sense, all scientific statements are inevitably quantitative, although the counting and measurement may be only implicit. Even a case study, for example, makes some implicitly quantitative statements: When we describe certain patterns of behavior that are characteristic of the person or community studied, then we are essentially saying that this pattern occurs *frequently* and in *many* kinds of situations. Often, moreover, we select a particular case for investigation because it is deviant or extreme—in other words, because it possesses a particular trait or set of traits to a *greater* or *lesser* extent than other comparable cases.

Similarly, scientific statements generally have a qualitative component, although it too is not always made explicit. Propositions that state the relationship between two quantitative variables apply only if certain background conditions obtain. For example, the nature of the relationship between social pressure and conforming behavior depends on the nature of the situation in which the pressure is exerted, on the cultural context in which the behavior is observed, and on the motivational set that the person brings to the experience. In principle, these background factors can all be stated in quantitative terms, but in practice we rely—and must rely—on qualitative statements of the limit-

ing conditions of our propositions. In empirical work, it is easy to ignore these conditions, because they are built—as constants—into the natural situations we select for study, or the experimental situations we create in the laboratory. Nevertheless, these qualitative factors have an important bearing on our findings and—no matter how strong our predilection for quantitative statements—must be taken into account.

Thus, the crucial issue that cuts across the quantitative-qualitative dichotomy is whether, in any given study, the investigator is collecting the data appropriate to his particular purpose, and whether he is drawing conclusions that his particular data entitle him to draw. Those of us who use quantitative methods must always ask ourselves whether what we are counting and measuring is meaningfully related to the phenomena we are interested in, or whether we are just counting and measuring whatever happens to be readily available. The primitive, simple-minded quantifier, who decides what to measure on the basis of the instruments he has at hand, with little regard to the problems he hopes to illuminate, is certainly deserving of criticism. His culpability is compounded if he mistakes his irrelevant or partial index for the real thing. A hypothetical example of such a misuse of the quantitative approach would be an investigator who is interested in the level and determinants of "religiosity" in modern society, who takes church attendance as his sole index of religiosity because it can be assessed easily and reliably, and who then proceeds to draw broad conclusions about religiosity on the basis of these data. Critics of a quantitative approach would undoubtedly point to a study of this kind as a good illustration of the foolishness of trying to investigate religiosity with quantitative methods. What is wrong with this study, however, is not that the investigator used quantitative methods, but that he based his quantitative indices on the wrong data—or at least on an insufficient range of data—and that the data he did collect did not entitle him to draw the conclusions that he proceeded to draw. These errors are not inherent in the use of quantitative analysis as such.

For qualitative research, too, the basic question is whether the investigator has obtained the data appropriate to the conclusions that he wishes to draw. The primary danger here is not that he will focus on irrelevant data that can be easily measured, but that he will focus entirely on positive instances. Qualitative research does not have as many built-in safeguards against the tendency to find what one is looking for. An investigator eager to illustrate a particular phenomenon may find many cases that are consistent with his analysis, while missing many other cases that might be inconsistent with it. But, again, this fallacy is not inherent in the use of qualitative analysis as such. In qualitative as well as in quantiative analysis, the question is: What kinds of data does the investigator use for his analysis, and does he take the nature of these data into account in drawing his conclusions? Specifically, in evaluating qualitative re-

search, we would ask whether the investigator has deliberately looked for disconfirming cases as much as for confirming ones—whether he has arranged his observations in such a way that contrary findings have an equal opportunity to emerge.

A second aspect of the rigor vs. vigor controversy on which I wish to comment briefly is the apposition between a holistic and an elementalistic approach to human behavior. One component of the "vigor" syndrome (though by no means a universal one) is the insistence on dealing with "the whole man," in contradistinction to segmental, elementalistic approaches which only deal with parts of man and destroy his basic unity. To my mind, this is a false issue, because I do not regard the whole man as the business of social science. He is the proper business of poetry, philosophy, religion, everyday interpersonal relations and, to a certain extent, of psychotherapy. But the task of the social and behavioral sciences is to dissect and analyze the behavoir of men and societies, to break it down in terms of theoretical constructs and genotypical formulations, and in this way to increase our systematic understanding of it. It is in the nature of social research to deal with parts of man, artificially separated out from the richness and unity of the total personality. The uniqueness and wholeness of the individual disappears at the hand of the generalizing social scientist, just as the poet's nature disappears at the hand of the natural scientist.

The holistic vs. elementalistic dichotomy is sometimes used to distinguish between psychoanalysis and academic psychology. But certainly psychoanalysis does not deal wtih the whole man. Among the greatest contributions of psychoanalytic theory are the different ways in which it slices man—and, until recently, the theory has not even concerned itself with some of these slices. It is precisely because of the fact that psychoanalysis is a scientific—hence not a holistic—system, that it is not a proper philosophy of life and cannot substitute for an adequate metaphysics and ethics. Man's ultimate meaning, his place in the universe, the nature of good and evil—all of these must be examined outside the scientific realm.

While social science is, by definition, concerned with parts, the picture of man that emerges from its theory and research must be *consistent* with the nature of man and society as they manifest themselves in their phenotypical wholeness. For example, a theory that cannot encompass such human and societal characteristics as self-sacrifice and love, or power and murder, is obviously of limited value. Another and related issue is whether the units of analysis that are used in a given theory and research program are appropriate to the problem with which they are intended to deal. This, it seems to me, is the real issue behind much of the holistic vs. elementalistic controversy. For example, I would question the use of a stimulus-response model for social psychology, not because it is too elementalistic, but because its units—which may be quite appro-

priate for the study of conditioning—are not equally appropriate to the complexities of social interaction. I would favor such concepts as social role and self-presentation—though they too are analytical and segmental—because I regard them as more appropriate to the level of analysis at which the social psychologist operates.

In sum, I have maintained that the rigor vs. vigor controversy and its various sub-controversies—such as those involving the relative virtues of quantitative vs. qualitative or holistic vs. elementalistic approaches—tend to focus on false issues. The real questions—stated in their most general terms—on which the evaluation of a given line of social research hinges, are how *systematic* and how *imaginative it is*. These questions are often masked by the rigor vs. vigor debate, because they sound deceptively similar to that dichotomy, while in fact they cut across it.

To be systematic is not the same as to be rigorous—in the sense of using experimental methods and precise, quantitative measures. Systematic work in social science refers to an organized and disciplined way of thinking about social behavior and of moving back and forth between conceptualization and evidence. Systematicness is not linked to a particular set of methods, although it does imply an awareness of the limitations of the methods one is using and of the conclusions that can properly be drawn from them. It is a relevant criterion in the evaluation of any social research, regardless of its form, and within each research tradition there may be wide variations in how adequately this criterion is met. Thus, qualitative, clinical work may be quite systematic, even though it is not rigorous. Conversely, quantitative, experimental work—though highly rigorous—may be quite unsystematic, if it involves the mere accumulation of empirical data without any attempt to relate these to efforts at conceptualizing social behavior.

The criterion of imaginativeness also cuts across the rigor vs. vigor dichotomy. The tendency to credit the "looser," clinical approaches with profound insights and to equate the "cold," statistical approaches with arid recapitulations of the obvious is a romantic notion that does not always conform to reality. The use of naturalistic methods, which attempt to capture the richness and the real-life flavor of the phenomenon under investigation, does not guarantee that the research will be imaginative. Conversely, quantitative and experimental work of the most rigorous kind may well be highly imaginative and imbued with a creative spark.

The real issue in the evaluation of social research revolves, in short, around the thinking and the imagination that it represents. There is no substitute for good thinking and good imagination, no matter what methods we use. And no method automatically insures us of either one of these.

EXPERIMENTAL RESEARCH IN SOCIAL PSYCHOLOGY

In distinguishing between the different purposes of different types of research, I pointed out that naturalistic research is particularly appropriate for exploring the dimensions of a problem and developing hypotheses, while experimental research is particularly appropriate for testing hypotheses about the functional relationships between different variables. This is, essentially, the usual distinction between research undertaken for purposes of discovery and research undertaken for purposes of verification. The latter clearly calls for a greater degree of rigor—highly structured designs, quantitative methods, and statistical analysis. By and large, I would agree with this formulation, but I would argue that—at least in social psychology—the distinction between discovery and verification is not as sharp as we sometimes maintain. In line with this position, I would like to propose a somewhat different perspective for viewing the functions and contributions of experimental research in this field.

The central features of research design in experimental studies are based on the assumption that we are engaged in efforts to verify general propositions. To this end, we are concerned about random assignment of our subjects to experimental conditions, about providing appropriate controls for our crucial comparisons, and about eliminating alternative hypotheses. In short, we set up our studies in such a way that they will allow us to verify propositions according to the usual standards of experimental method. If experimental research is to maximize its potential contributions—which, as I shall point out shortly, I regard as very considerable—then its practitioners must follow these procedures faithfully. The value of this work rests on our playing the experimental game according to its rules. And, as long as we do, I feel that we are entitled to the privilege of talking the experimental game—of using the language of testing and verification adopted from the natural sciences.

Let us not, however, deceive ourselves about the status of what we have found when we complete an experiment in social psychology. I would maintain that the findings of social-psychological experiments—even if they have been replicated a dozen times—can hardly be thought of as experimentally verified and established laws of nature. I assume that it is reasonable to view experimental findings in the natural sciences—and perhaps even in certain areas within psychology—in these terms. In social psychology, however, the gap between the laboratory and the real world is so great that one is hardly justified in the conclusion that what has been established in the laboratory constitutes a verified fact about nature. This is not because the laboratory situation is unreal, but because it has its own reality. The characteristics of the experimental situation in which we put our hypotheses to the test are related in ways that are largely unknown to the characteristics of the situations to which we hope to

generalize our findings. First of all, the laboratory situation as such has unique characteristics of its own, which are only partly understood—although there has recently been a healthy concern with exploring the social psychology of the psychological experiment. In view of the special characteristics of the laboratory situation, it may well be that the whole array of findings based on laboratory studies is applicable only to behavior in the laboratory—and in a subset of other social situations that are similar along certain crucial dimensions. In addition to the unique characteristics of laboratory situations in general, the situation created for any given experiment or series of experiments has special characteristics of its own. As I pointed out earlier, in discussing the role of qualitative factors, in setting up an experiment we devise certain background conditions that remain constant throughout. It may well be that the relationships found hold only in situations that share some of these background conditions. Similarly, in any given experiment we operationalize our independent variable in a particular way and we measure our dependent variable in a particular way. We know that the particular ways in which our variables are operationalized and measured often make a difference in the relationships obtained. It is very difficult, therefore, to have any reasonable assurance that relationships found in the laboratory apply to the wide range of real-life situations, characterized by different background conditions and different manifestations of the independent and dependent variables.

In view of the idiosyncratic and unrepresentative nature of experimental situations in social psychology, and in view of our limited knowledge of the dynamics of those situations, we cannot reasonably equate the confirmation of an experimental hypothesis in the laboratory with the verification of a general principle in nature. Our ability to draw such conclusions may be enhanced as we explore a particular relationship in a wide variety of settings—both experimental and natural, and as we learn about the special characteristics of the laboratory situation. At least as of now, however, our field has certainly not achieved this stage in its development.

Does that mean, then, that our experimental work is really *just* a game, that we are merely playing scientist without contributing anything substantial to an understanding of the general principles of social behavoir? My answer is an emphatic no. In my view, experimental research can make enormously important contributions to social-psychological knowledge. These contributions, however, take the form of providing *unique inputs into systematic thinking about social-psychological processes*, rather than of establishing laws about social behavior.

I regard systematic thinking about man and society as the central and essential task of the social scientist. I would not insist that it is his *only* legitimate task. He may also, for example, apply his skills to the solution of certain practi-

cal problems. Or he may devote himself to the collection and processing of various kinds of social data (such as demographic or public opinion data) because of their historical interest or policy relevance, rather than because of their theoretical significance. But systematic thinking—continually confronting all manner of evidence and aiming for the development and refinement of general propositions—is the social scientist's task *par excellence*. It is in this context that I would want to evaluate the contributions of experimental methods, and it is in this context that I see a unique role for the experimental approach. There are at least four ways in which experimental research can feed into the process of thinking about social behavior more effectively than any other type of method:

(1) The requirement to translate our concepts into experimental operations imposes a discipline on our thinking that might otherwise be lacking. As long as we remain at the level of manipulating words, we can gloss over certain conceptual difficulties and avoid the necessity of really resolving certain ambiguities in our thinking. Once we attempt, however, to specify the conditions necessary for testing our propositions, to create laboratory situations encompassing these conditions, and to manipulate our variables through concrete operations, we begin to discover ambiguities that had remained unnoticed and we are forced to face difficulties that we had been avoiding. The necessity of devising an experiment forces us to commit ourselves—to state clearly what our concepts mean and to pin down precisely what relationships we expect.

(2) Experiments offer us an opportunity to observe causal relationships, which can usually be inferred only indirectly and tenuously from other types of evidence. An experiment cannot, of course—as I have already stressed—indicate the generality of the relationship found. It may, however, provide a very important input into our thinking by showing that the causal relationship between two variables *can* be in the particular direction found, at least under certain circumstances. This kind of information is useful in identifying the dimensions on which our conceptual efforts ought to focus and in suggesting lines of thought that are likely to be more or less productive to pursue.

(3) Experiments provide operating models of the social-psychological systems that we are interested in exploring. They allow us to study certain processes in situations that we have deliberately created and that, therefore, have some advantages over the real-life situations to which we ultimately hope to generalize. The experimental situations are simpler, being stripped—insofar as possible—of extraneous variables and historical complications; they are situations whose histories and dimensions are more fully known to the investigator; and they are—at least to a limited extent—subject to the investigator's control. It is thus possible to observe the operation of the specific variables of interest in a detailed and relatively uncontaminated fashion. It is also possible

to extend our range of observations to hypothetical situations that do not exist and have never occurred in real life. Along with its advantages, the stylized and artificial character of the experimental situation also contains certain disadvantages—notably the difficulty in generalization. When taken in conjunction with observations from real life, however, the observations of a hypothetical model in action provide unique inputs into our thinking about social processes and their potentialities for change.

(4) If theoretical thinking is to remain productive, it cannot feed entirely upon itself, but requires periodic stimulation by new inputs from outside sources. For the social scientist, his empirical observations constitute the major source of such inputs. Every type of observation has something unique to offer as a stimulus to new thinking. One of the unique contributions of experimental observations is that they often derive from novel, atypical situations. Another unique contribution is the possibility of accumulating findings from a series of systematically interrelated experiments, which together point a new direction. It is interesting to note that unanticipated experimental findings—which are not too desirable from the point of view of verifying propositions—are particularly useful inputs into new thinking. Unanticipated findings call the investigator's attention to variables he had not thought of before and suggest interpretations and qualifications that he had not considered.

These are some of the special ways, then, in which experimental work feeds into the process of thinking about social behavior and that make the experimental method such an important tool in social research. It is on their embeddedness in this longer conceptual process, rather than on their direct contribution to the body of verified laws about social behavior, that the significance of experimental studies rests. *An experimental finding, at least in our field, cannot very meaningfully stand by itself.* Its contribution to knowledge hinges on the conceptual thinking that has produced it and into which it is subsequently fed back.

There are several implications to the present view that an experimental finding cannot stand by itself. The most obvious implication, with which no theoretically oriented experimenter would disagree, is that *findings from any single experiment* cannot stand by themselves. Contributions are of necessity cumulative; only as a series of experiments—either by the same investigator, or by different investigators working on related problems—build upon each other, can we begin to formulate meaningful conclusions. There are different points of view about what constitutes a good research program, likely to produce a cumulative effect. Some experimenters prefer to narrow in on their problem, using a variety of experimental situations that focus on related issues, in the hope that the nature of the phenomenon they are exploring will gradually become clarified. Others prefer to use a single experimental situation, systematically varying

all the variables that are—for theoretical or eimpirical reasons—potentially significant, in the hope that they can thus pin down the whole array of factors controlling the phenomenon under investigation. Both types of experimental programs may provide useful inputs into thinking about social behavior, though each has different strengths and weaknesses. Thus, the "narrowing-in" approach provides a better basis for assessing the generality of the phenomenon, while the "varying all the variables" approach is more useful in systematically identifying the dimensions that ought to be considered.

A second implication of the present view, with which some experimental social psychologists may not concur, is that *findings from experimental research* cannot stand by themselves. Because of the gap between laboratory situations and the range of real-life situations to which we want to be relevant, we cannot base our general propositions on the findings from experimental research alone. Our thinking must be informed by data from a wide variety of sources—population surveys and correlational research, participant observation and community studies, analysis of documents and of individual cases, historical studies and ethnographic reports. This is not to say that every investigator must work in all of these traditions, or even in more than one. There is no reason why a social psychologist who is trained in experimental work, is good at it, and enjoys it, should not devote himself to laboratory studies. He should, however, be aware of the limitations of experimental work and to its place within the larger context of systematic thinking about social behavior. Moreover, if he is well trained, he should at least be able to draw on observations from other sources as he contemplates his experiment—both in the process of defining his problem and developing his hypotheses before he begins his experiment, and in the process of checking the generality of his findings and exploring what they mean in action after the experiment is over.

Finally, the present view implies that *empirical facts* in our field—especially, but not exclusively, those based on experimental research—cannot stand by themselves. It is not the facts that constitute a contribution, but what is done with them. It may, of course, happen that an experimenter does little or nothing with his findings himself, but they become the raw material for the conceptual work of others. In any event, the ultimate value of experimental findings depend on the quality of the thinking in which they are embedded. It follows that a study that is procedurally clean and well-designed, but unrelated to a serious conceptual problem, is less valuable than a messier study that forms part of a systematic process of thinking about social behavior. In other words, even though rigor is one of the unique and central contributions of experimental method, rigor as an end in itself is self-defeating. In the final analysis, no method can susbstitute for intelligence and imagination.

ABSTRACT

Many dubious issues turn on one or another aspect of what might be called the "rigor vs. vigor" controversy. They are generally "false issues" of questionable validity, such as naturalistic vs. experimental, psychological vs. sociological, quantitative vs. qualitative, and holistic or molar vs. elementalistic or reductive approaches.

The real issue is whether a particular approach is appropriate to the question the investigator is trying to answer, and the even more complicated matter of whether the question is worth asking.

Rigor, the stressing of hard facts through the use of exact methods, is contrasted with vigor, stressing methods that will capture real-life flavor in all its fullness, and this dichotomy is used to illustrate that the choice of a proper method depends on the nature of the problem under investigation, the stage of development of the research area, the specific problem on which the work is being done, and the preferred style of the investigator. The value of the research depends upon how systematic and imaginative the approach is, for no experimental finding in our field can stand meaningfully by itself, apart from the conceptual thinking and related theory, or apart from cumulative supportive evidence.

RESUMEN

En la controversia del "rigor" versus "vigor," se presentan muchos aspectos problemáticos que son de dudosa validez y generalmente falsos. Estos enfoques aparecen en pares que se han dado en llamar "naturalista versus experimental," "psicológico versus sociológico," "cuantitativo versus cualitativo," y "el todo versus las partes."

Un aspecto realmente importante es constatar si un determinado enfoque está relacionado o no con la pregunta a la cual el investigador trata de buscar respuesta. Y lo que es más siginificativo aún, es si realmente vale la pena efectuar la pregunta.

El *rigor,* con un énfasis en la constatación de hechos por medio del uso de métodos exactos, es comparado con el *vigor,* que pone un énfasis en métodos que recogen el sabor de la vida en su totalidad. Esta dicotomía se usa para ilustrar el hecho de que la elección de un método apropiado depende de la naturaleza del problema que se va a investigar, de la etapa de desarrollo en que se encuentre la correspondiente área de investigación, el problema específico sobre el cual se hará la investigación y el estilo preferido del investigador.

El valor de la investigación dependerá directamente de cuán sistemático e imaginativo sea el enfoque. Aparte del pensamiento conceptual y una teoría

relacionada, o de una evidencia acumulada que lo respalde, ningún resultado de investigación en nuestro campo tiene sentido en forma aislada.

RESUMO

Na controvérsia entre "rigor" versus "vigor," muitos aspectos se confrontam. Estas discussões são geralmente de carácter duvidoso, tal como naturalista versus experimental, psicológico versus sociológico, quantitativo versus qualitativo e todo versus partes.

O problema real é se a abordagem específica que está sendo usada, é adequada para responder a pergunta proposta. Ainda mais importante é saber se a pergunta em sí, merece consideração.

Rigor, ou a ênfase na constatação de fatos através de métodos exatos, é comparado com *vigor,* ou a ênfase em métodos que se propõe a capturar o sabor da vida na sua totalidade. Esta dicotomia é usada para ilustrar o fato de que a escolha de um método apropriado depende do problema a ser pesquisado, do nível de desenvolvimento do campo de pesquisa em sí, e o estilo preferido pelo investigador.

O valor da pesquisa dependerá de quão sistemática e imaginativa for a abordagem, pois nenhuma evidência experimental em nosso campo poderá ser significativa em sí mesmo, aparte do pensamento conceitual, existentes teorias e dados relacionados.

TRANSFERENCIA E INHIBICION RETROACTIVA CON BILINGUES[1]

ROBERT K. YOUNG
University of Texas at Austin, U.S.A.

El patrón de conducta del inmigrante que llega a los Estados Unidos se caracteriza por sus deficiencias en el manejo del idioma y por lograr éxito muy limitado en la sociedad. El hijo del inmigrante ha asimilado la cultura y el idioma y se desempeña mejor que su padre. El problema ha captado la atención del sociólogo. Este patrón de conducta presenta en la segunda generación un rechazo hacia la cultura y el idioma de origen del padre y una persistencia en aprender Inglés.

Sin embargo, un grupo que no sigue este patrón y se ha resistido a asimilar la cultura de los Estados Unidos es el de los inmigrantes de México (e incidentalmente los de Puerto Rico). Por diversas razones el inmigrante mejicano ha retenido la cultura y la lengua de su tierra en forma mucho mayor que cualquier otro grupo en los Estados Unidos. Una de las más importantes es la cercanía de Mexico, que permite una continua participación en la cultura de México además de la mantención de los lazos familiares.

En Texas existe una población de 1,500,000 personas con apellidos españoles. El hecho de que un gran porcentaje de ellos habla Español y que un porcentaje semejante es, desgraciadamente, analfabeto en esta lengua repercute en el tipo de investigación que nos proponemos efectuar.

El fenómeno tiene su origen en el sistema educacional del estado de Texas y en la escala de valores de las personas relacionadas con las escuelas públicas. El deseo de asimilar todo inmigrante a la cultura de los Estados Unidos, juega un rol decisivo en el manejo de las escuelas públicas. Sin embargo, la lengua hispana ha subsistido en los hogares mejicanos de los estados colindantes con México y los educadores de esos estados ya han comenzado a reconocerlo. Sintetizando, se cree que quizás comenzar desde el primer grado hasta el fin de la escuela secundaria a enseñar todo en Inglés al niño de origen mejicano, no es el mejor sistema de educación. En estas condiciones el niño de ascendencia mejicana llega a la escuela y se encuentra de inmediato retrasado e incapaz de

[1] For a recent comment on this type of unanticipated use, see J. Graciarena, "Algunas Consideraciones Sobre la Cooperación Internacional y el Desarrollo Reciente de la Investigación Sociológica" in *Revista Latinoamericana de Sociologia*, Vol. 65, No. 2, p. 240. For additional descriptions of archives, see Ralph L. Bisco, "Social Science Data Archives: a Review of Developments." in *American Political Science Review*, Vol. LX, No. 1, March, 1966, pp. 93–109, and *Information sur les Sciences Sociales*, Vol. V, Mars, 1966, pp. 110–125.

competir con sus compañeros de habla inglesa. La experiencia escolar no puede proporcionar muchas satisfacciones si se desarrolla en una lengua que no es familiar.

El hecho de que hablar Español dentro del recinto escolar sea motivo de castigo y que al niño no se le enseñe ni a leer ni a escribir en su lengua natal es obviamente un tremendo desperdicio de potencial humano. Al mismo tiempo que al niño de habla hispana se le prohibe hablar Español en la escuela, los Estados Unidos necesita en forma desesperada representantes del Gobierno o comerciales en América Central y en Sud América que sean fluídos en Español. Los programas escolares que se han desarrollado en este sentido muestran que nos estamos autofrustrando. En todo caso, los problemas que el niño bilingue presenta en la escuela, el hogar y la comunidad han despertado gran interés. El hecho de que no se hace uso de los mejores y más productivos métodos en la educación de bilingues, es un hecho que se está reconociendo lentamente.

Con estos antecedentes hemos llevado a cabo nuestro estudio sobre el bilingue. Hasta ahora nada se ha hecho por investigar los efectos del bilinguismo en al aprendizaje.

Nuestro marco de investigación es uno de los más antiguos en psicología. Ebbenghaus comenzó sus estudios sobre el aprendizaje humano y la retención alrededor de 1880. Esta es una área que ha atraído considerablemente la atención y el esfuerzo constante de muchos investigadores desde esa época hasta ahora. Así, el área de la retención puede ser considerada clásica al compararla con otras áreas en psicología ya que se ha hecho investigación en ella durante casi 100 años. En general el investigador busca las variables que influyen en el aprendizaje de la retención de listas de material verbal tales como palabras o sílabas sin sentido. El investigador quiere saber por qué algunas palabras se recuerdan mejor que otras, por qué algunas son más fáciles de retener que otras y por qué una lista fué más fácil de aprender una vez que se siguió el mismo metodo de aprendizaje empleado en la otra lista. Fué dentro de este marco que condujimos nuestra investigación sobre bilinguismo. Queríamos encontrar una respuesta a simples preguntas como: ¿Ejerce el aprendizaje de una lista de palabras en un determinado idioma alguna influencia en el aprendizaje de la misma lista cuando ésta se traduce a un segundo idioma? ¿El aprendizaje de una lista en un idioma determinado, interfiere o facilita el aprendizaje de la misma lista en un segundo idioma? ¿Una vez que se ha aprendido una segunda lista de palabras en un segundo idioma, es más difícil recordar la lista en el primer idioma? Para dar respuesta a estas preguntas, seguimos el siguiente procedimiento. Un estudio inicial, conducido por Joel Saegert consistió en requerir a 48 alumnos de la escuela secundaria que aprendieran dos listas conteniendo 12 sustantivos comunes cada una. El método de aprendizaje para todas las listas, fué el de anticipación de series. Substancialmente la tarea del sujeto

al aprender una lista en serie es recordar un conjunto de ítems, en un orden constante. Así la tarea es como aprender el alfabeto o una canción.

En este caso se utilizó un grupo de sustantivos inconexos y la tarea de los sujetos consistió en aprender a recordar cada sustantivo antes de que éste fuera presentado. Los sustantivos se presentaron por medio de una grabadora debido a que los alumnos podían hablar Español pero no leerlo. La mitad de los sujetos aprendió una lista en Español y después una lista en Inglés mientras que la otra mitad aprendió primero una lista en Inglés y luego una lista en Español. En ambos casos los resultados fueron los mismos. (No mencionaremos esta variable otra vez durante el resto del experimento).

Se usó tres grupos de transferencia básicos. Al primero, el grupo de control, se le dió una segunda lista de palabras qué además de estar en otro idioma no tenía ninguna relación con la primera lista. El segundo grupo, con el cual se empleó la técnica de usar el mismo orden en las palabras, aprendió una lista y luego una segunda en que los ítems estaban en el mismo orden y eran una simple traducción de los primeros. Por ejemplo: MESA, CASA, SILLA, CALLE, etc. y luego: TABLE, HOUSE, CHAIR, STREET, etc. El tercer grupo, aprendió las listas con los sustantivos al azar. Los ítems eran los mismos, pero al traducirlos se alteró el orden de la lista. Por ejemplo, MESA, CASA, SILLA, CALLE, etc., fué luego aprendida: CHAIR, TABLE, STREET, HOUSE, etc.

Los resultados de este experimento fueron directos. Al ser ambos grupos comparados con el grupo de control hubo transferencia positiva en el grupo en que el mismo orden se mantuvo y una leve transferencia negativa en el grupo en que se usó un orden al azar. Después de 12 intentos el grupo de control logró aprender la segunda lista. El grupo en que se mantuvo el orden efectuó solo 6 tentativas, lo que significa un 50% de transferencia positiva, mientras que el grupo que usó orden al azar efectuó 14 tentativas, o sea un 15% de transferencia negativa. Con este experimento hemos demostrado que el aprender asociaciones en una lengua puede influir la formación de asociaciones en otra. Se obtiene transferencia positiva al traducir un ítem sin alterar su ubicación en la lista, pero si ésta es alterada, obtenemos transferencia negativa.

Un aspecto incidental interesante de estos resultados es que al usar este mismo diseño de investigación con monolingues, pero en lugar de traducir los ítems, hacer uso de sinónimos, no se presenta transferencia positiva ni negativa al aprender la segunda lista.

Nuestro laboratorio está ahora investigando por qué hay transferencia de Inglés a Español, o de Español a Inglés pero no de Inglés a Inglés.

El segundo experimento fué conducido por Arthur Webber. En oposición al primero, aquí se empleó la transferencia de paradigmas que han sido usados extensamente en investigaciones tradicionales de aprendizaje con monolingues.

Esto paradigmas que en varias oportunidades han producido transferencia, los hemos adaptado para usarlos con bilingues. El procedimiento usado en este segundo experimento, en contraste con el primero, fué aprendizaje de pares asociados. En este tipo de aprendizaje el individuo aprende varios sustantivos pareados de manera que cuando ve el primer miembro del par (el estímulo) trata de recordar el segundo miembro (la respuesta). Cada estímulo se muestra durante dos segundos y luego el estímulo y la respuesta por dos segundos. Después de esto se muestra un estímulo diferente y otra vez, después de dos segundos, la respuesta, también durante dos segundos. En contraste con el procedimiento de aprendizaje en serie los pares no se presentan en la misma secuencia, sino más bien en secuencias diferentes de manera que el sujeto los aprende como pares más bien que como secuencias.

En el segundo experimento se usó 40 alumnos universitarios bilingues en Español e Inglés. Al igual que en el primer experimento se usaron 3 condiciones. La primera, llamada "Misma condición," presentaba una primera lista de pares que luego eran traducidos y aprendidos como segunda lista de pares. Por ejemplo, si la primera lista presentaba MESA - CASA, CALLE - SILLA, la segunda presentaba TABLE - HOUSE, STREET - CHAIR. La segunda condición llamada "Re-pareada," presentaba en la segunda lista ítems traducidos pero no consistentes con la primera. Por ejemplo, si los pares MESA - CASA y CALLE - SILLA aparecían en la primera lista, en la segunda aparecerían como TABLE - CHAIR y STREET - HOUSE. La tercera condición fué una condición de control en la cual los reactivos de la segunda lista no guardaban ninguna relación con los de la primera. Al igual que en el primer experimento la mitad de los sujetos fueron expuestos primero al Español y luego al Inglés y la otra mitad vice-versa. Estos sujetos sabían leer en Español.

Los resultados de este experimento fueron en parte semejantes a los del primer experimento. Bajo la "Misma condición" los sujetos efectuaron 5.50 tentativas antes de aprender las segundas listas mientras que los sujetos bajo la condición de control tuvieron que efectuar 7.50 intentos antes de aprenderla — o sea una transferencia positiva de aproximadamente un 27%. Por otra parte se encontró transferencia negativa solamente en el grupo Español-Inglés del grupo Re-pareado. Este grupo debió efectuar 9.20 tentativas antes de aprender la lista. Al comparar esta cifra con 7.00 en el grupo de control obtenemos un 30% de transferencia negativa. Sin embargo los grupos Inglés-Español del grupo Re-pareado y el grupo de control debieron efectuar 12.50 tentativas para aprender sus listas.

Debido a experiencias anteriores con monolingues esperábamos encontrar transferencia positiva en el grupo bajo la "Misma condición," y asi lo comprobaron nuestros resultados. Además se predecía que se encontraría transferencia negativa en el grupo Re-pareado. Esto ocurrió solo en parte. La ausen-

cia de transferencia negativa en el grupo Inglés-Español sugirió que ocurriría transferencia negativa solamente en aquellas condiciones en que la segunda lista estuviera en la lengua dominante del sujeto que la aprendía. Casi todos los bilingues que aprendieron las listas de este experimento se desempeñaban mejor en Inglés que en Español. Suponiendo que las asociaciones se efectúan en la lengua dominante, esperábamos que existiera interferencia al aprender una lista en la lengua secundaria (mientras las asociaciones se efectuaban en la otra lengua) y aprender después una segunda con los mismos ítems de la primera, en la lengua dominante. Por ejemplo, si la primera lista aprendida contenía pares como MESA—CASA, CALLE—SILLA, el sujeto cuya lengua dominante es el Inglés formaría asociaciones TABLE—HOUSE y STREET —CHAIR. Si después se le pide que aprenda una lista de pares en Inglés como TABLE—CHAIR y STREET—HOUSE, las asociaciones aprendidas anteriormente van a interferir con las ahora presentadas y por ejemplo, STREET va a ser asociada con CHAIR, que fué lo que aprendió en la primera lista, mas bien que con HOUSE. De manera que la asociación STREET—CHAIR va a interferir con la asociación del par STREET—HOUSE. El aprendizaje de esta lista sería por lo tanto más difícil de aprender que la lista de control, que no tendría ningún tipo de interferencia.

Por otro lado, los sujetos cuya lengua dominante es el Inglés y aprenden una lista primero en Inglés van a formar asociaciones en Inglés aún al aprender la segunda lista.

Durante el aprendizaje de la segunda lista, el sujeto encuentra que traducir del Español al Inglés para formar asociaciones produce una interferencia en las asociaciones ya formada en la primera lista. El resultado de esto es que los sujetos prefieren formar sus asociaciones en Español lo que dificulta el aprendizaje y lo hace más lento. En este experimento los pares re-pareados y los pares de control estaban en la misma lista. Debido a que los sujetos tienen dificultades en formar asociaciones en Inglés entre algunos de los pares, puede que adopten la técnica de formar todas las asociaciones en Español. Como consecuencia de esto, tanto el aprendizaje de la lista de control como el de la lista re-pareadad ha resultado más lento y más difícil.

De modo que parece que la lengua dominante es una variable de suma importancia en el estudio de la transferencia en los bilinques.

El tercer experimento que condujimos en nuestro laboratorio estaba especialmente relacionado con la lengua dominante como variable. En este experimento conducido por Isabelle Navar, otra vez usamos 40 alumnos universitarios bilingues, ninguno de los cuales había participado en el experimento anterior. La lengua dominante de la primera mitad de los sujetos era el Inglés, la de la otra mitad, el Español. Dentro de cada uno de los grupos la mitad de los sujetos aprendió una lista de pares asociados en Español y

luego en Inglés. La segunda mitad aprendió una lista de pares asociados en Inglés y luego en Español. La relación de transferencia fué similar a la usada en el experimento anterior cuando encontramos una transferencia positiva uniforme. En este experimento buscamos el paradigma de transferencia re-pareada una vez más. Asi cada sujeto aprendió una lista y luego una segunda lista en la cual los pares de la primera fueron traducidos, pero las respuestas no iban acompañados de sus estímulos.

Las respuestas fueron pareadas con estímulos que habían sido traducidos de la lista anterior pero con respuestas que no habían sido pareadas anteriormente. Una vez aprendida la primera lista, se asignó un tiempo de 20 minutos para aprender la segunda. A los sujetos que la aprendieron en menos de 20 minutos, se les dió una tarea de cancelar símbolos hasta completar el tiempo requerido, de manera que después de 20 minutos todos comenzaron a aprender la primera lista otra vez.

Hubo bastante inhibición retroactiva u olvido. El grupo de control efectuó 1.80 tentativas hasta aprender la primera lista de nuevo y el grupo Re-pareado, 3.80, o sea mas del doble del tiempo empleado por el grupo de control.

Otras medidas de respuesta dan los mismos resultados; el grupo "Re-pareado" olvidó la primera lista mucho más facilmente que el grupo de control. En contraste con el estudio anterior, ni la lengua dominante ni el orden de las lenguas se presentó como variable. Es decir se produjo olvido ya la lengua dominante del sujeto fuera el Inglés o el Español, o ya aprendiera primero una lista en Español o en Inglés. Dedujimos por lo tanto que la lengua dominante no es una variable.

¿A qué conclusiones nos llevan estos experimentos? La principal conclusión es que los idiomas de un bilingue no pueden ser considerados dos conjuntos de respuestas independientes sino mas bien respuestas recíprocas en que las asociaciones formadas en una lengua influyen enormemente en las asociaciones formadas en la segunda. Bajo ciertas circunstancias la formación de asociaciones en la segunda lengua se verá facilitada por las asociaciones de la primera mientras que en otras, la formación de una segunda lista de asociaciones será interferida por las asociaciones ya existentes en la primera lengua. Por último, los datos indican que la formación de asociaciones en la segunda lengua está relacionada al olvido de las asociaciones ya existentes en la primera legua.

El mecanismo preciso en relación a la interferencia y al olvido entre los idiomas no es fácil de identificar. Será muy sencillo decir que cada palabra en Inglés está relacionada a su traducción en Español. Sin embargo, otros estudios han mostrado que muchos bilingues fluídos en ambas lenguas no pueden traducir facilmente de un idioma a otro.

Bajo tales circunstancias, sería difícil concluir que palabras en Inglés

están asociadas a las palabras en Español como pares de sinónimos. Sin embargo las asociaciones en una lengua parecen interferir con las asociaciones en la segunda al extremo de ser olvidadas. Las implicaciones de esta investigación en términos de aplicación práctica son inmensas y en algunos casos totalmente obvias, como por ejemplo el hecho que la enseñanza de todas o la mayoría de las materias de la escuela primaria, tales como historia o aritmética se facilitaría si se enseñaran en la lengua dominante del alumno. De este modo, si el estudiante aprendiese un segundo idioma o quisiera lograr más fluidez en él, tendría la información a su alcance en cualquiera de los dos idiomas. Por otra parte, si el material se presenta en un idioma secundario que el alumno no domina lo suficiente, ni siquiera para su comprensión, será imposible que lo aprenda. Esto no significa que el niño, por ejemplo, sea incapaz de aprender historia, sino que en primer lugar no posee las herramientas básicas del idioma para aprender el material.

RESUMEN

El problema que presentan los niños bilingues de ascendencia mejicana en Texas ha sido objeto de interesantes investigaciones. Este artículo trata de responder algunas preguntas relacionadas con el aprendizaje y la retención en conexión con la adquisición y asociación de la primera y la segunda lengua.

Se efectuaron tres estudios diferentes presentando dos listas de palabras sueltas y de pares de palabras primero en Inglés y luego en Español y viceversa a sujetos bilingues. La muestra del primer experimento la constituyeron 48 alumnos bilingues de la escuela secundaria que podían leer y hablar en Inglés pero que solo hablaban y comprendían Español. El segundo experimento incluyó 40 alumnos universitarios bilingues cuya lengua dominante era el Inglés pero que leían y hablaban Español. El tercer experimento contó con una muestra de 40 alumnos universitarios bilingues. La lengua dominante de la mitad de ellos era el Inglés y la de la segunda mitad, el Español.

La principal conclusión que se deriva de estos experimentos es que los idiomas del bilingue no constituyen dos sistemas de respuestas independientes uno del otro y que por el contrario las asociaciones que se forman en una lengua son vehementemente influenciadas por las asociaciones formadas en la otra. Las implicaciones que esta conclusión tiene en los programas de la escuela primaria para niños bilingues son obvias.

ABSTRACT

The Mexican American bilingual in Texas is an interesting subject for research. The present study attempts to answer some questions concerning

learning and retention in connection with first and second language acquisition and association.

Three studies were conducted presenting the subjects two lists of isolated words of pairs of nouns, first in English and then Spanish, and vice-versa. The subjects of the first experiment were 48 high school students who could speak and read in English but could only speak and understand Spanish. The second experiment included 40 Spanish English bilingual college students whose dominant language was English but could also read and speak Spanish. The third experiment was carried out with a group of 40 college students, all of them bilinguals but half of them had English as their dominant language and the other half had Spanish.

The main conclusion drawn from the experiments was that the languages of a bilingual are not two independent set of responses and that the associations formed in one language are strongly influenced by the associations formed in the other. The implications of this conclusion towards the elementary school programs for bilingual children are obvious.

RESUMO

O problema especial que caracteriza as crianças de descendência mexicana nascidas no Texas, por serem bilingües, tem sido objeto de interessantes pesquisas. Este trabalho representa uma tentativa de responder certas perguntas relacionadas com aprendizagem e retenção, relativamente à aquisição e associação da primeira e segunda língua.

Três estudos foram realizados apresentando-se a sujeitos bilingües duas listas de palavras individuais e de palavras em pares, primeiro em inglês, em seguida em espanhol e vice-versa. A amostra do primeiro estudo era composta de 48 alunos de nível secundário, que liam e falavam inglês mas sòmente falavam e compreendiam espanhol. O segundo estudo contou com 40 estudantes universitários, cuja língua dominante era o inglês, mas que eram capazes de falar e escrever o espanhol. O terceiro estudo foi também realizado com estudantes universitários bilingües. A língua dominante da metade dêles era o inglês, e da outra metade, o espanhol.

A conclusão mais importante proveniente dêstes estudos é que os idiomas do bilingüe não constituem sistemas de resposta independentes um do outro, mas antes, que as associações formadas na outra língua. As implicações desta influenciadas pelas associações formadas na outra língua. As implicações desta conclusão para os programas das escolas primárias para bilingües são óbvias.

[1]La presente investigación fué en parte realizada gracias a la donación número GB-3629 que la National Science Foundation hiciera al autor y a David T. Hakes y este trabajo presentado en el Primer Congreso Nacional de Psicología auspiciado por la Universidad Veracruzana y la Sociedad Mexicana de Psicología, del 20 al 23 de Marzo de 1967 en Xalapa, Veracruz, México.

THE INTERNATIONAL DATA LIBRARY AND REFERENCE SERVICE

DAVID NASATIR
University of California, Berkeley, U.S.A.

This note is an attempt to describe one example of a new and growing movement in the social sciences: the data archive. It has long been recognized that while many sociological studies are not initially planned to validate the hypotheses of later investigators, they may nevertheless serve such a purpose. In addition, they provide information useful in the design and interpretation of other studies.[2] Therefore, in 1958, the Institute for International Studies at the University of California (Berkeley) and the Survey Research Center of the University began to collaborate in the development of a campus facility to serve the needs of local scholars in this regard. Sample survey data conducted by scholars and, in some cases, commercial polling organizations was sought for the archive. The data came at the specific request of individual users—in the form of punched cards and the various documents associated with the conduct of the original studies. Proceeding through informal channels, the Survey Research Center began to acquire the basic, raw data collected by various scholars in their own work. It then made this data available to other scholars in order that it might be analyzed further and in combination with other studies. These analyses were usually carried out with a perspective distinct from that of the original authors and produced many new conclusions.

As the amount and diversity of the information stored at the Survey Research Center increased, so did the number and variety of users of the facility. Graduate students utilized the data archive for writing dissertations; local and foreign scholars began to use it for supplementing their work. In 1961, a social science data archive was formally established as part of the Survey Research Center including in its holdings the many studies executed

[1] For a recent comment on this type of unanticipated use, see J. Graciarena, "Algunas Consideraciones Sobre la Cooperación Internacional y el Desarrollo Reciente de la Investigación Sociológica" in *Revista Latinoamericana de Sociología*, Vol. 65, No. 2, p. 240. For additional descriptions of archives, see Ralph L. Bisco, "Social Science Data Archives: a Review of Developments." in *American Political Science Review*, Vol. LX, No. 1, March, 1966, pp. 93–109, and *Information sur les Sciences Sociales*, Vol. V, Mars, 1966, pp. 110–125.

[2] Students and scholars from such diverse countries as Argentina, Brazil, Chile, Colombia, France, Germany, India, Italy, Japan, Korea, Mexico, Taiwan and Venezuela have used the archive in their own research and as an aid in teaching the techniques of survey analysis to students.

by the Center itself, other organizations on the Berkeley campus, and materials from individual scholars at Berkeley and other campuses.

The archive continued to solicit contributions of data from various scholars without developing a specific substantive focus until 1963. During that year, Professor S. M. Lipset (the Director of the Institute for International Studies) and Professor C. Y. Glock (Director of the Survey Research Center) agreed to sponsor a joint program of data acquisition emphasizing comparative studies with special emphasis upon the developing nations, and formally announced the establishment of the International Data Library and Reference Service (IDL & RS). Since that time the IDL & RS has continued to develop as a general purpose service archive but has given special emphasis to that part of its total holdings dealing with sample survey data from the developing countries, particularly those in America and Asia.

In 1966, the University of California, in collaboration with the Chinese University of Hong Kong established the Social Survey Research Centre in Hong Kong. The operations of this Centre are coordinated with those of the Survey Research Center in Berkeley, and data gathered by the Hong Kong Centre will become part of the holdings of the International Data Library and Reference Service.

Until late in 1964, the IDL & RS focused its activities almost exclusively on acquiring data. While involved in these activities personnel of the IDL & RS were also able to become more familiar with the problems and the prospects for obtaining survey data from the developing countries: locating potential sources, making contact with them, establishing personal relations, and making the legal and technical arrangements necessary before data and supporting documentation could be shipped to Berkeley. Toward the end of this year, sufficient preliminary work had been accomplished along these lines to begin to turn away from the problems of acquisition to emphasize those associated with use of the data Archive and dissemination of its holdings to scholars.

Since there were, at that time, well over one hundred different studies from countries in Africa, Latin America and Asia (each study carried out independently and according to different customs and procedures), the problems of making them useful for secondary analysis were considerable. Funds to aid in this endeavor were therefore, sought, and obtained, from the National Science Foundation of the American Government. These funds permitted the staff in the Archive to devote increasing attention to problems of translation of codebooks, checking and verifying the correspondence of data with codebooks, obtaining supporting documentation and similar activities.

The holdings of the Library have grown since that time, and the problems of organizing and maintaining the Archives have grown with them. In order to keep up with the increasingly complex demands of the users, therefore, the

Archive has shifted, more and more, toward a reliance upon electronic computers for certain kinds of basic operations. The inventory of holdings, for example, is maintained by computer. Similarly, in reproducing specific studies (or part of studies) for users, or in carrying out statistical computations that they have requested, the computer is utilized extensively. Funds have recently been obtained to foster this activity and to develop, if possible, a direct connection of the Berkeley Archive with other archives in order to enable users of the IDL & RS to gain rapid access to data from those archives which may be either more general or more specialized in their holdings.

The costs of utilizing the IDL & RS are kept at a minimum through the indirect support provided by the University of California and the National Science Foundation funds used for program development. Thus, only the direct costs of reproducing and shipping the desired data or documents are charged to the users. While these costs may change, they are, at the present, approximately $6.00 per thousand IBM cards and $.05 per page for the reproduction of codebooks and other documents. The costs of making statistical calculations vary, of course, with the complexity of the calculations desired and the number of respondents and variables involved. Estimates of the costs for any specific computation may be obtained through correspondence. When the number of respondents and the number of variables in an analysis are large, it is frequently most economical to utilize the computer for making such computations (even though the computations themselves may be relatively simple) because of its great speed.

More detailed information including a schedule of fees and a list of studies currently held in the Archive can be obtained by requesting them from

Data Librarian
International Data Library and Reference Service
Survey Research Center
University of California
2220 Piedmont Avenue
Berkeley, California 94720

CRITICA

Elaboración de Tests Psicológicos por Dorothy Adkins Woods, versión castellana, México, Editorial F. Trillas, S.A., 1965, 157 págs.

La obra publicada tiene una serie de características que hacen de ella un texto de gran utilidad para el psicólogo de habla castellana. En efecto, en el campo de las pruebas mentales y de la medición psicológica existen pocos textos

traducidos y este viene a llenar un vacío evidente al referirse, en particular, a la construcción de instrumentos objetivos para la evaluación psicológica y educativa. Reúne a su favor la claridad de quien como Dorothy Adkins ha dedicado su vida a este tema. Su conocimiento es tan profundo que la permite una de las cosas más difíciles de lograr en la literatura técnica: sintetizar en muy pocas páginas los asuntos básicos imprescindibles para la construcción de pruebas y ésto sin perder información, con solidez y sin descuidar la sutileza. Por su sencillez el libro está al alcance de una amplia gama de lectores, quienes tal vez se amedrentarían frente a obras más extensas, analíticas y técnicas como las de Lindquist o Guilford, para citar solo dos, podrán leer la obra de la Dra. Adkins rapidamente y adquirir, al mismo tiempo, conocimientos precisos. Tambien sacarán provecho de ella los psicólogos más avezados que hallarán soluciones concretas y abundantes ejemplos de tipo práctico no desprovistos de originalidad.

La elección de la obra para su traducción nos parece excelente. No podemos decir lo mismo, en cambio, sobre la traducción. Damos por descontado las múltiples dificultades que debe resolver quien quiere vertir de modo fidedigno una obra técnica de psicología y sabemos el esfuerzo que representa todo aporte en este sentido. Pero creemos que hay en esta traducción muchos términos que producirán confusión en el lector. No deseo hacer una crítica exhaustiva de todos los términos y giros que en mi opinión han sido mal elegidos. Es necesario, sin embargo, señalar que es poco acertada la traducción de "item" por *reactivo*. No vemos por qué si se acepta *test*, palabra inglesa de origen latino, no puede aceptarse *ítem*. En todo caso podría haberse traducido por elemento, problema o pregunta, que es en definitiva lo que representa un ítem sea o no formulado en forma interrogativa. Además se mezcla a veces con la palabra *estímulo* y desde las primeras páginas se presenta sin ninguna nota que aclare el significado. Dado que existen trabajos originales de autores de habla castellana de prestigio de M. Yela, H. Rimoldi, F. del Olmo y otros que no han dudado en usar la palabra *ítem*, nos parece innecesario pretender ahora introducir una nueva palabra. Igualmente consideramos inadecuado usar el vocablo de *métrica* psicológica cuando tenemos incluído en el acerbo cultural propio el de *medición*. Tampoco nos parece felíz la elección de *calificación* para "score." En realidad calificación es sinónimo de *nota* y para "score" se suele usar *puntaje* o *puntuación* (si se quiere ser más purista). También existen términos que pueden dar lugar a confusión de conceptos. Asi "range" debe traducirse por *amplitud* o *extensión*. Jamás debe traducirse por *rango* pues éste representa un lugar en una escala jerárquica y es la traducción del término inglés "rank" (Este error es bastante usual en las traducciones). En el libro de Atkinds no coinciden algunas veces los vocablos del índice analítico con los del texto y se presentan innovaciones totalmente superfluas. ¿Por qué llamar *diseminación* a

la dispersión o variabilidad?. "Standard deviation" suele traducirse clasicamente por *desviación típica*, pero si se quiere usar la palabra inglesa "standard," castellanizada debe escribirse con *e* inicial, acento y sin la *d* final (estándar). En fín, para terminar, consideramos que en obras de este tipo tan técnico la traducción debería ser revisada muy cuidadosamente y por lo menos tratar de no innovar en los que ya se ha resuelto bastante satisfactoriamente. Para las traducciones estadísticas la consulta de las obras clásicas traducidas anteriormente y del glosario en castellano del "Dictionary of Statistical Terms" de Kendall, que tiene la garantía de haber sido aprobado por UNESCO, es una verdadera necesidad. En pocas palabras, la obra es muy buena y de gran interés para los psicólogos de lengua castellana, pero sería de desear que en adelante se tomaran las precauciones necesarias para lograr versiones más cuidadosas.

Nuria Cortada de Kohan
Universidad de Buenos Aires, Argentina

BOOKS RECEIVED LIBROS RECIBIDOS LIVROS RECEBIDOS

Rojas Boccalandro, Efraim, *Teste Não Verbal da Inteligência G-36,* Vetor Editôra Psico-pedagógica Ltda., São Paulo Brazil.

Hall, Calvin S. and Lindzey, Gardner, *Teorias da Personalidade,* Herder Editôra Livraria Ltda., São Paulo, Brazil.

Allport, Gordon W., *Personalidade,* Herder Editôra Livraria Ltda., São Paulo, Brazil.

Allport, Gordon W., *Desenvolvimento da Personalidade,* Herder Editora Livraria Ltda., São Paulo, Brazil.

Giberti, Eva and Escardó F., *Hospitalismo.*

Ausubel, D., *The Fern and the Tiki,* Holt, Rinehart and Winston, Inc., New York, N.Y.

Ausubel, D., *Maori Youth,* Holt, Rinehart and Winston, Inc., New York, N.Y.

Hanson, J. and Brembeck, C., *Education and the Development of Nations,* Holt, Rinehart and Winston, Inc., New York, N.Y.

Wolcott, H., *A Kwakiutl Village and School,* Holt, Rinehart and Winston, Inc., New York, N.Y.

Warren, R., *Education in Rebhausen—A German Village,* Holt, Rinehart and Winston, Inc., New York, N.Y.

Gay, J. and Cole, Michael, *The New Mathematics and an Old Culture,* Holt, Rinehart and Winston, Inc., New York, N.Y.