



Revista Latinoamericana de Psicología

ISSN: 0120-0534

direccion.rlp@konradlorenz.edu.co

Fundación Universitaria Konrad Lorenz

Colombia

León, Federico R.

Control del "error de reciprocidad" en la evaluación del desempeño docente

Revista Latinoamericana de Psicología, vol. 10, núm. 3, 1978, pp. 351 - 361

Fundación Universitaria Konrad Lorenz

Bogotá, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=80510304>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

CONTROL DEL "ERROR DE RECIPROCIDAD" EN LA EVALUACION DEL DESEMPEÑO DOCENTE¹

FEDERICO R. LEÓN *

Escuela de Administración de Negocios para Graduados

A few days after midterm exams, thirty graduate Business students taking an Organizational Behavior course rated the teaching performance of the instructor (an assistant professor); fifteen of them did it before they knew their grades in the exam and the other fifteen after they knew their grades. In the "before" sample the correlation between ratings and grades was .02 while in the "after" sample the correlation was .42 ($p < .05$, one-tailed). Simultaneously, the same students rated the teaching performance of a full professor of Accounting, but here the students of the first sample knew their grades while those of the second sample ignored them; the correlations between ratings and grades in the Accounting course were .50 ($p < .05$, one-tailed) in the first sample and .05 in the second sample. It was concluded that the validity of student evaluations of faculty can be significantly improved by obtaining them before grades are made public.

Es creciente la utilización del juicio de los alumnos en la evaluación del desempeño del profesor universitario, pero tal método

¹ Konrad Fischer propuso el tema de esta investigación; Hans Frank aceptó servir de conejillo de Indias junto con el autor; Rosa Piscoya de Pérez Costa obtuvo las evaluaciones del desempeño docente con la colaboración de Enrique Valdez y Martin Scurrah, ingeniándose para contribuir a la investigación sin que se revele la identidad de los alumnos que suministraron las evaluaciones; con la autorización de Luis Flores, Violeta Pastor ayudó a que los alumnos recibieran sus notas en el momento oportuno; y Elsa Ames realizó los análisis estadísticos. A todos ellos va el agradecimiento del autor.

* Dirección: Faculty of Management Studies, University of Toronto, 246 Bloor Street West, Toronto, Ontario, Canadá M5S 1V4.

no está libre de problemas. Entre éstos parece tenerse una tendencia del alumno a efectuar la evaluación influido por la nota que ha recibido del profesor: mientras peor sea la nota obtenida por el alumno, es probable que más crítico sea éste al evaluar al profesor. Schuh y Crivelli (1973) propusieron el término "error de animadversión" para describir este fenómeno. La elección del término, sin embargo, puede no haber sido la más afortunada, pues sugiere que, al evaluar al profesor, sólo los alumnos de bajo rendimiento son influidos por la nota recibida; en algunos círculos se cree que los alumnos favorecidos con las mejores calificaciones también son influidos por el conocimiento de la nota —aunque no negativa sino positivamente, por supuesto— al hacer las evaluaciones. Ambos errores, el de animadversión y el otro, podrían ser incluidos entonces dentro del concepto más amplio de "error de reciprocidad", por el cual el alumno califica al profesor en función de la calificación dada a él por éste. Pero ¿existen realmente estos fenómenos o no pasan éstos de ser meras suposiciones de los profesores?

Schuh y Crivelli (1973) han demostrado que existe un sesgo en las evaluaciones que hacen los alumnos respecto al desempeño docente del profesor y que tal sesgo se define por el hecho de tener notas buenas o malas. A mitad de un curso de Relaciones Industriales estos investigadores encontraron una correlación significativa entre la evaluación del desempeño del profesor por parte de los alumnos y la nota que los alumnos indicaban haber recibido de él, e interpretaron este hallazgo como un resultado del "error de animadversión"; es decir, la correlación se habría debido a que los alumnos con malas notas, al conocerlas, reaccionaron agresivamente contra el profesor, evaluándolo negativamente.

En el estudio de Schuh y Crivelli, sin embargo, no se llegó a establecer que fuera el conocimiento de las notas lo que influyó en las evaluaciones. El de Schuh y Crivelli fue un estudio de campo de diseño corriente —no un experimento— y como tal deja a sus resultados abiertos a interpretaciones alternativas. Se podría conjeturar, por ejemplo, que algunos alumnos evaluaron negativamente al profesor porque éste no hizo el curso atractivo para ellos, y al mismo tiempo obtuvieron notas bajas debido a su escaso interés en el curso; éstas constituirían condiciones suficientes para crear una correlación positiva entre notas y evaluaciones sin que la correlación reflejara una relación causal de la forma "las notas —causan— las evaluaciones". Queda pues por establecer de modo riguroso si el conocimiento de la nota *per se* es lo que ejerce una influencia sobre la evaluación del desempeño docente del profesor.

El asunto es de interés "teórico" pero, como todo asunto de esta naturaleza, tiene importantes implicaciones prácticas; en este caso ellas son de índole administrativo y social. Si se estableciera la

existencia de una relación de causa-efecto entre notas y evaluaciones habría que concluir que los profesores, intencional o inintencionalmente, pueden maximizar la bondad de las evaluaciones de su desempeño docente merced a su control sobre el proceso de asignación de notas: mientras más altas sean las notas que asignen, mejores serán las evaluaciones que reciban por parte de los alumnos. Si esto fuera así, ello haría incomparables las evaluaciones que reciben los profesores más "exigentes" y las que obtienen los profesores más "generosos", pues a igualdad de otras condiciones, los últimos resultarían con evaluaciones superiores que no reflejarían una superioridad real en su desempeño docente. Debe parecer obvio que, en esta situación, sería administrativamente impropio y socialmente injusto usar las evaluaciones como elemento estándar de evaluación del desempeño académico del profesor sin haber introducido antes en ellas los ajustes necesarios.

En el experimento que se describe a continuación se intentó determinar si hay una relación de causa-efecto entre notas y evaluaciones y si esta fuente de invalidez en las evaluaciones puede ser controlada mediante el simple procedimiento de recoger las evaluaciones antes de entregar las notas.

METODO

Procedimiento

En la segunda sesión de clase posterior a la semana de exámenes de mitad de trimestre, los alumnos de una sección del Programa Magister en Administración de la Escuela de Administración de Negocios para Graduados de Lima, Perú, ESAN (XIV Promoción) recibieron cuestionarios de evaluación de cursos a razón de uno por alumno. Los cuestionarios se referían a dos de los 5½ cursos del programa trimestral. A media sección —elegida aproximadamente al azar— se le pidió evaluar el curso de Contabilidad y el desempeño del profesor principal a cargo del mismo, y a la otra mitad, el curso de Comportamiento Organizacional y el desempeño de un profesor auxiliar a cargo de una parte de él, de manera que mientras un alumno evaluaba a un profesor los dos vecinos de sus costados evaluaban al otro. Al hacerlo todavía no tenían información sobre las notas obtenidas en los exámenes de medio trimestre. Al terminar la clase, los alumnos recibieron las notas de Contabilidad y Comportamiento correspondientes a estos exámenes y, después de almorzar, entraron nuevamente a clase para encontrarse con un segundo pedido de evaluación. En este caso, a los alumnos que habían evaluado al profesor de Contabilidad (Comportamiento) en la mañana se les pidió evaluar al profesor de Comportamiento (Contabilidad), según el modelo

de la Tabla 1. Ninguno de los dos profesores se encontraba presente al recogerse las evaluaciones, y en cada período se aseguró a los alumnos que ningún profesor de la Escuela conocería la identidad de los informantes individualmente.

TABLA 1

*Diseño Experimental**Sección B del Programa Magister*

	Evalúa al profesor de Contabilidad (Datos A) N = 16	Evalúa al profesor de Comportamiento (Datos B) N = 15
Segunda sesión de clase en la mañana.		
Entrega de notas		
Primera sesión de clase en la tarde.	Evalúa al profesor de Comportamiento (Datos C) N = 15	Evalúa al profesor de Contabilidad (Datos D) N = 15

(N indica el número de casos utilizables en cada grupo. El número de casos para C no es 16 porque en la segunda fase de la encuesta un alumno del Grupo 1 comprendió mal las instrucciones y vició sus evaluaciones).

Datos

Dos clases de datos fueron analizados: las notas parciales y las evaluaciones. La nota parcial de Contabilidad, que representaba el 30% de la nota total del curso, provenía de un examen multifacético de mitad del curso que incluyó problemas de clasificación de cuentas, asientos contables, etc., y preguntas de tipo verdadero-falso. La información que sobre su rendimiento en el curso tenían hasta entonces los alumnos se limitaba a los resultados de dos pasos que contribuirían —junto con otros, posteriores— al 10% de la nota total. En el caso de Comportamiento, la nota parcial provenía de un examen "objetivo" de tipo elección múltiple que cubría los temas de siete sesiones que estuvieron a cargo del profesor auxiliar y tres sesiones iniciales que estuvieron a cargo de un profesor asociado; esta nota parcial representaba 40% de la nota total del curso. La información que tenían hasta entonces los alumnos sobre su rendimiento en el curso comprendía un puntaje asignado por el profesor asociado —el cual contribuiría el 10% de la nota total del curso— e información ambigua sobre los resultados de dos pasos para la parte del

profesor auxiliar; estos pasos —junto con otros pasos y dos trabajos de grupo— contribuirían a otro 10%.

La evaluación del desempeño docente de cada profesor se hizo en una escala de 1 (Malo) a 5 (Excelente) puntos en respuesta a la siguiente pregunta: "En forma integral, ¿cómo evaluaría al profesor de este curso?". Este ítem seguía a 28 ítems del cuestionario de evaluación de cursos que se referían al curso y al desempeño del profesor en varias áreas específicas.

RESULTADOS

Comparabilidad de las muestras

La división al azar de la sección en dos grupos produjo dos muestras comparables en notas parciales: ni en Contabilidad ($t = .14$, no significativo) ni en Comportamiento ($t = .98$, n.s.) existieron diferencias significativas entre los grupos 1 y 2 en lo que se refiere a notas promedio logradas en los exámenes de medio curso. Asimismo, las notas de Contabilidad, aunque fueron ligeramente más altas, resultaron comparables a las de Comportamiento, tanto en el Grupo 1 ($t = 1.83$, n.s.) como en el Grupo 2 ($t = 1.26$, n.s.).

Respecto a evaluaciones del desempeño docente, aquellas efectuadas por los alumnos antes de conocer las notas resultaron ligeramente más altas, pero no difieren significativamente en promedio de las que fueron efectuadas por los alumnos después de conocer las notas; esto se aplica tanto al caso del profesor de Contabilidad ($t = 1.32$, n.s.) como al de Comportamiento ($t = 3.40$, n.s.). En cambio las diferencias en evaluación entre los dos profesores sí resultaron significativas; el profesor de Contabilidad fue evaluado más positivamente que el de Comportamiento, tanto en el Grupo 1 ($t = 7.79$, $p < .01$, dos colas) como el Grupo 2 ($t = 4.77$, $p < .01$, dos colas). Estas diferencias son científicamente interesantes, pues permiten verificar en dos niveles de desempeño docente la existencia del error de reciprocidad.

Relaciones notas-evaluaciones

Los resultados centrales del estudio conciernen a las relaciones entre notas y evaluaciones para cada juego de datos. La Figura 1 muestra —en dos de estos juegos— las relaciones que existen entre las notas asignadas por el profesor de Contabilidad a los alumnos y las evaluaciones que él recibe de éstos. Recuérdese que el juego A corresponde al Grupo 1, donde las evaluaciones fueron hechas por los alumnos antes que éstos conocieran su nota; el juego D corresponde al Grupo 2, donde las evaluaciones fueron efectuadas por los alumnos después de conocer la nota. La relación entre notas y eva-

Aquaforest TIFF Junction Evaluation

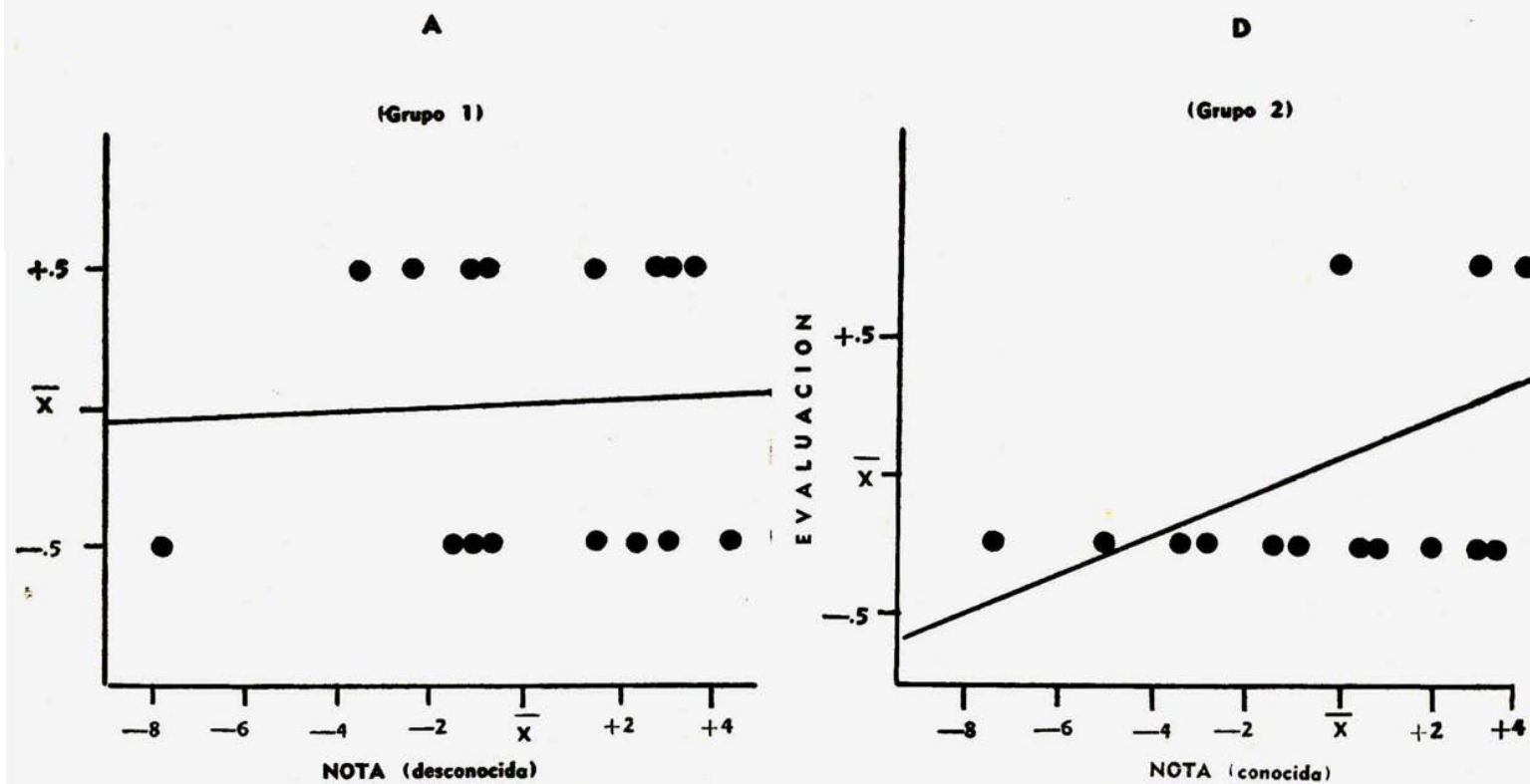


FIGURA 1. Regresión de evaluaciones sobre notas para los juegos de datos A y D (Contabilidad). En cada caso las escalas representan desviaciones del promedio correspondiente al juego de datos que se describe.

luaciones resultó trivial en el primer caso ($r = .05$, no significativa) mientras que en el segundo caso resultó considerable ($r = .50$, $p < .05$, una cola).

La Figura 2 presenta las distribuciones bivariadas para el curso de Comportamiento, donde se repite el fenómeno observado en el de Contabilidad: las evaluaciones no guardan relación alguna con las notas en el Grupo 2, donde los alumnos ignoraban sus notas (caso B; $r = .02$, n.s.) pero sí en el Grupo 1, donde las conocían (caso C; $r = .42$, $p < .05$, una cola). El diseño experimental usado en este estudio permite hacer inferencias causales a partir de tales datos.

El antecedente común en los casos A y B es el hecho de que los alumnos produjeran las evaluaciones *antes* de conocer sus notas; como consecuente se tiene el hecho de que las evaluaciones no guardan relación alguna con el nivel de rendimiento académico del alumno. En los casos C y D el antecedente común es el hecho de que los alumnos produjeran las evaluaciones *después* de conocer sus notas; aquí como consecuente se tiene el hecho de que quien logra notas altas tiende a calificar alto al profesor, mientras que el que logra notas bajas tiende a calificarlo bajo. La asociación entre notas y evaluaciones se observa, pues, sólo cuando el alumno conoce su nota. Esto permite concluir que el conocimiento de la nota *influye directamente* sobre la manera como el alumno evalúa al profesor.

DISCUSION

El experimento llevado a cabo muestra que las evaluaciones del desempeño docente que efectúan los alumnos tienden a estar contaminadas por las notas que éstos reciben del profesor y conocen, más no por las que reciben y desconocen*. Estos resultados plantean un problema y al mismo tiempo sugieren la manera de resolverlo. En el presente estudio, los profesores de Contabilidad y Comportamiento fueron igualmente exigentes (o igualmente generosos) al asignar notas parciales a los alumnos; de allí que *no* se pueda afirmar que uno fue evaluado por los alumnos más positivamente que el otro en

* ¿Son generalizables estos hallazgos a todas las situaciones donde un profesor universitario es evaluado por sus alumnos? Brunswik (1956), con su preocupación por la "validez ecológica" de los experimentos, afirmaría que para justificar esta generalización se tendría que haber encontrado los resultados en una muestra al azar del universo de tales situaciones, en lugar de dos muestras que resultaron cómodas para efectos de la investigación que acaba de ser descrita. Pero Campbell y Stanley (1966, pp. 16-18, 32-36) probablemente preferirían poner el énfasis en el hecho de que ya hay varias evidencias consistentes en favor de la hipótesis del presente estudio, la cual ha sido puesta a prueba rigurosamente sólo en las muestras de ESAN, pero también ha sobrevivido una prueba menos rigurosa en una muestra tan "diferente" como la de Schuh y Crivelli (1973).

Aquaforest TIFF Junction Evaluation

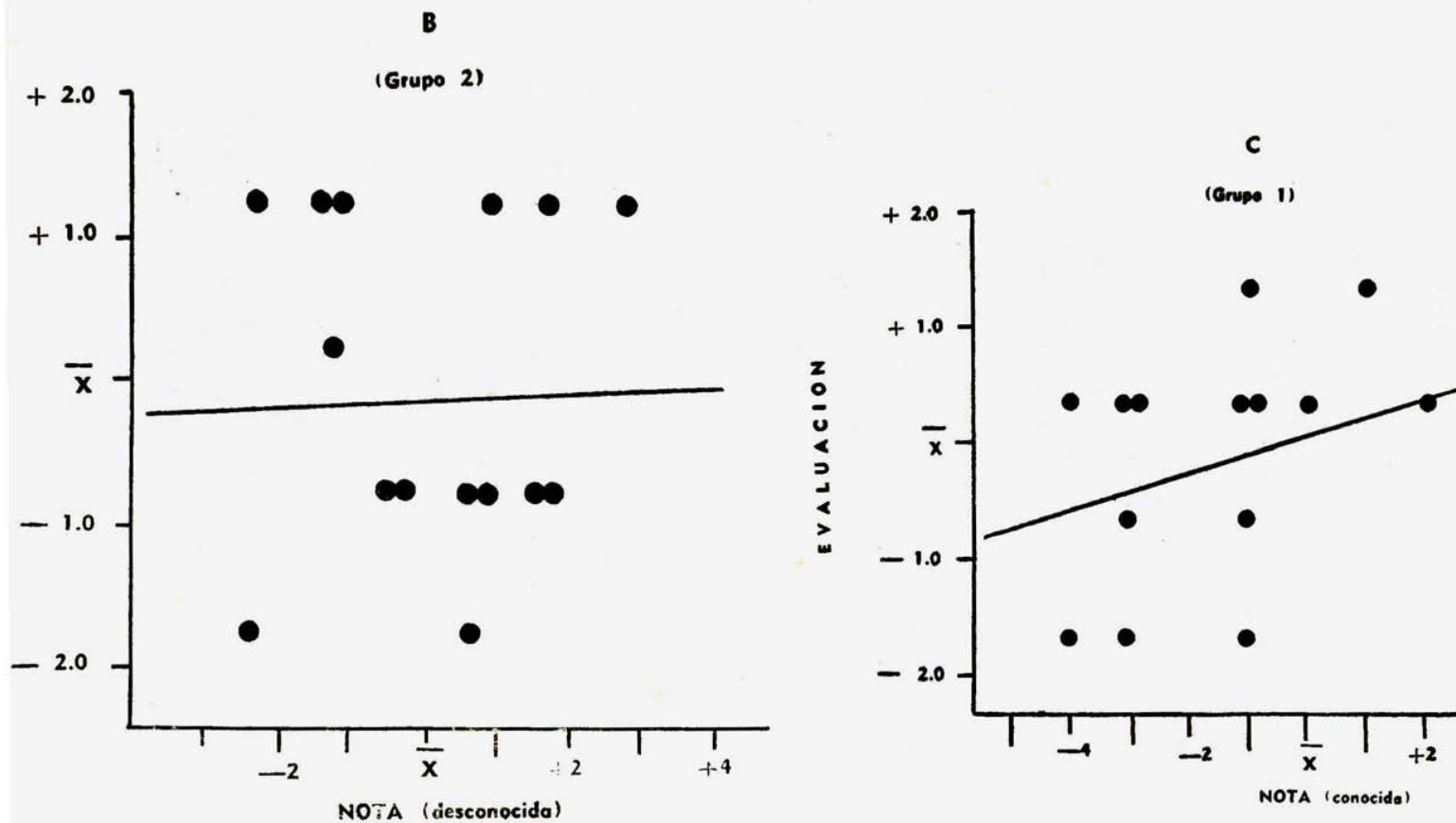


FIGURA 2. Regresión de evaluaciones sobre notas para los juegos de datos B y C (Comportamiento). En cada caso las escalas representan desviaciones del promedio correspondiente al juego de datos que se describe.

virtud del error de reciprocidad. En todo ámbito académico, sin embargo, existen profesores exigentes y profesores generosos entre los cuales sí hay diferencias significativas en cuanto a nivel de notas asignadas. Es en estos casos donde el error de reciprocidad puede causar diferencias espúreas en la evaluación del desempeño docente de los profesores, *favoreciendo a los profesores generosos* y perjudicando a los profesores exigentes. Los resultados del estudio indican que para eliminar este peligro se cuenta con el simple expediente de recoger las evaluaciones antes de la entrega de notas. El hecho de que las correlaciones entre notas y evaluaciones fueran de alrededor de cero cuando las últimas se recogieron antes de entregar las notas permite afirmar que este procedimiento basta para ejercer el control necesario sobre la fuente de invalidez del juicio estudiantil que se analizó aquí.

Nótese, sin embargo, que este procedimiento no resuelve todos los problemas involucrados en la evaluación del desempeño del profesor por parte de los alumnos y que, a su vez, plantea otros problemas. En primer lugar, se debe considerar que, si se quiere controlar el error de reciprocidad, no sólo es necesario recoger las evaluaciones como paso previo a la entrega de notas sino que esto tiene que producirse hacia mitad del curso, pues al término del mismo el juicio del alumno ya estaría contaminado por la nota parcial. De establecerse el procedimiento de recoger las evaluaciones a mitad de curso, sin embargo, los administradores universitarios se privarían de la oportunidad de obtener información relativa al desempeño del profesor en la segunda parte, a no ser que decidan recoger varios juegos de evaluaciones para el mismo curso. La segunda parte de un curso puede exigir habilidades y/o conocimientos distintos de los requeridos por la primera parte, y si esto es así, el desempeño del profesor en ella también debería ser evaluado. De otro lado, el procedimiento de recoger las evaluaciones sólo a mitad de curso podría estimular a los profesores a concentrar su esfuerzo en la primera parte, descuidando la segunda, y ésta es una razón más para sugerir que no se deje de evaluar su desempeño en la segunda parte.

En segundo lugar, se debe tener en cuenta que el procedimiento de recoger las evaluaciones antes de entregar las notas puede ayudar a eliminar el error "objetivo" de reciprocidad que se ha analizado en este estudio, pero no necesariamente el error "subjetivo" de reciprocidad que puede cometer el alumno al evaluar al profesor en función de la nota que él *cre*e que éste le va a asignar, lo cual es una creencia eventualmente manipulable por el profesor. Estudios hechos por Bausell y Magoon (1972), Schwab (1975), y Centra y Linn (1976) revelan una correlación significativa entre la evaluación del curso o del desempeño del profesor por el alumno, y la nota esperada por éste en el curso. Afortunadamente, los mismos estudios indican que parece ser pequeño el impacto que sobre la evaluación del

desempeño del profesor tiene *la nota esperada* por el alumno; bastante menor, en efecto, que el impacto de la nota entregada.

Por último, no habría que olvidar que el temor a represalias del profesor es un factor que puede influir en la evaluación que de su desempeño hace el alumno. Si el alumno no confía en la promesa que se suele hacer en el sentido de que las evaluaciones bajo ninguna circunstancia provocarán represalias, podría ser mejor solicitar las evaluaciones contra entrega de la nota final, pues así se reduciría —en la percepción del alumno— la probabilidad de que *el profesor* cometa errores de reciprocidad al asignar notas. Aún bajo las condiciones de confianza que permitirían recoger las evaluaciones a mitad de curso habría que enfatizar ante el alumno que las notas, si bien no han sido entregadas todavía al alumno, lo han sido ya por el profesor a las oficinas de administración; además, por supuesto, se tendría que remarcar que el profesor nunca sabría quiénes son los autores de las buenas evaluaciones y quiénes lo son de las malas.

RESUMEN

Pocos días después de los exámenes de medio trimestre, treinta estudiantes graduados en Administración que tomaban un curso de Comportamiento Organizacional evaluaron el desempeño docente del profesor (un profesor auxiliar); quince de ellos lo hicieron antes de conocer sus notas en el examen y los otros quince después de conocerlas. En la primera muestra la correlación entre notas y evaluaciones fue .02 mientras que en la segunda muestra la correlación fue .42 ($p < .05$, una cola). Simultáneamente, los mismos alumnos evaluaron el desempeño docente de un profesor principal a cargo de un curso de Contabilidad, pero aquí los estudiantes de la primera muestra conocían sus notas mientras que los de la segunda las ignoraban; las correlaciones entre notas y evaluaciones en el caso del curso de Contabilidad fueron .50 ($p < .05$, una cola) en la primera muestra y .05 en la segunda. Se concluyó que se puede mejorar significativamente la validez de las evaluaciones del desempeño docente obteniéndolas antes de dar a conocer las notas.

REFERENCIAS

- Bausell, R. B., y Magoon, J. Expected grade in a course, grade point average and student ratings of the instructor. *Educational and Psychological Measurement*, 1972, 32, 1013-1023.
- Brunswik, E. *Perception and the representative design of psychological experiments*. Berkeley: University of California Press, 1956.
- Campbell, D. T., y Stanley, J. C. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1966.
- Centra, J. A., y Linn, R. L. Student points of view in ratings of college instruction. *Educational and Psychological Measurement*, 1976, 36, 695-703.
- Schuh, A. J., y Crivelli, M. A. Animadversion error in student evaluations of faculty teaching effectiveness. *Journal of Applied Psychology*, 1973, 58, 259-260.
- Schwab, D. P. Course and student characteristic correlates of the Course Evaluation Instrument. *Journal of Applied Psychology* 1975, 60, 742-747.