

UNIVERSIDAD IBEROAMERICANA



IMPACTO DE LA DESEABILIDAD SOCIAL EN LAS EVALUACIONES DE
CAPACITACION

T E S I S

Que para obtener el grado de
DOCTOR EN INVESTIGACIÓN PSICOLÓGICA

P r e s e n t a

FRANCISCO JAVIER ENRÍQUEZ JIMÉNEZ

DIRECTOR:

DRA. ALEJANDRA DEL CARMEN DOMÍNGUEZ ESPINOSA

LECTORES

DR. CARLOS HÉCTOR DORANTES RODRÍGUEZ

DR. JESUS FELIPE URIBE PRADO

México, D. F.

2010

Contenido

RESUMEN.....	3
INTRODUCCIÓN.....	4
CAPITULO I EVALUACIÓN DE LA CAPACITACIÓN Y SU BENEFICIOS	8
CAPITULO II INSTRUMENTOS DE EVALUACION Y DESEABILIDAD SOCIAL....	16
CAPITULO III MÉTODO	31
Planteamiento del Problema	31
Hipótesis	32
Diseño.....	33
Participantes	33
Instrumentos	34
Procedimiento.....	35
Tratamiento Estadístico	36
CAPITULO IV RESULTADOS	38
CAPITULO V DISCUSION Y CONCLUSIONES.....	51
REFERENCIAS	61
ANEXO A	74
ANEXO B	75
ANEXO C	76
ANEXO D	77
ANEXO E.....	78

RESUMEN

El propósito del presente estudio fue el obtener evidencia del efecto de las instrucciones Socialmente Deseables en las evaluaciones de la capacitación. Se exploró el sesgo de respuesta sobre la evaluación del desempeño de los instructores (usando el Cuestionario de Evaluación de la Capacitación, CEC), causado o ligado al fenómeno de Deseabilidad Social (DS). Se consideraron dos tipos de instrucciones (Induciendo Respuestas Socialmente Deseables vs. Estándar) y dos tipos de instructores (interno vs. externo). El CEC cuenta con un índice adecuado de consistencia interna ($\alpha = .92$), una validez factorial que explican en su conjunto el 72.61% de varianza. Como medición de DS se utilizaron las escalas de Marlowe y Crowne (MCSDS) y el Inventario Equilibrado de Respuesta Deseable (BIDR) con una consistencia interna de $\alpha = .768$ y $\alpha = .804$, respectivamente. Los análisis de correlación entre el CEC y las escalas de DS mostraron asociaciones bajas pero significativas (MCSDS, $r = .235$, $p < .01$; BIDR, $r = .360$, $p < .01$). El CEC y BIDR bajo instrucciones inducidas vs instructor interno se correlacionan en un .484. El BIDR mostró una capacidad predictiva del 11% en el CEC. A pesar de tratarse de una evaluación de un tercero, hay evidencia de que las personas puedan estar sesgadas en sus juicios y presentar distorsiones en las respuestas atribuibles a la DS.

Palabras clave: Deseabilidad Social, Instrucciones Socialmente Deseables, Evaluación de la Capacitación, Psicología Organizacional.

INTRODUCCIÓN

Dentro del área de acción de los psicólogos organizacionales se encuentra la planeación y organización de las actividades de capacitación para los empleados en las cuales se invierte grandes cantidades de recursos económicos y humanos. Los cursos de capacitación incorporan dentro de sus sistemas de control de calidad, evaluaciones acerca de las actividades y de las habilidades de los instructores, con el objetivo principal de comprobar si la capacitación ha sido efectiva para los propósitos planteados. Sin embargo, los instrumentos con los cuales se realizan dichas evaluaciones, en términos generales, no son sujetos de escrutinio ni se comprueban sus cualidades psicométricas, por lo que en ocasiones las estimaciones desprendidas de los mismos pueden no ser eficaces o certeras en cuanto a la información proporcionada.

En los procedimientos de evaluación de la capacitación, frecuentemente se comete el error de realizar las evaluaciones sin garantizar el anonimato, o la confidencialidad de los resultados, siendo que las evaluaciones de los empleados llegan directamente a los jefes inmediatos. Esta situación para los respondientes, en ocasiones, es interpretada como una situación amenazante ya que ven “comprometida” su situación laboral a partir de lo que puedan o no responder, por lo que pueden distorsionar intencionalmente lo que se les pregunta, de tal modo que los resultados obtenidos pueden presentar una gran cantidad de error de estimación. A partir de las distorsiones intencionales sobre las evaluaciones, un instructor puede o no resultar favorecido; se le puede continuar contratando, darle más grupos para capacitar, etc. a pesar de no merecerlo. Por otra parte es posible que las respuestas a los cuestionarios de evaluación de la capacitación sean distorsionadas de modo

negativo, y pocos son los estudios que dan evidencia de este fenómeno. Este hecho debe despertar el interés de los profesionales de la Psicología del Trabajo y las Organizaciones que tienen responsabilidades en la evaluación y seguimiento de los programas de capacitación.

Gran parte de la literatura científica ha proporcionado evidencia acerca de las principales fuentes de distorsión como lo son la aquiescencia, las respuestas extremas y en particular la DS (Ray, 2001; Thomas et al., 1979). Esta última fuente de invalidación hace referencia a cuando las personas tienden a mostrarse a sí mismas de acuerdo a las normas sociales establecidas de una manera exagerada o poco habitual. La Deseabilidad Social (DS) implica sobre dimensionar las actitudes y comportamientos socialmente deseables y reducir los comportamientos y actitudes socialmente desaprobadas o menos deseables. La DS se evidencia cuando el sujeto escoge una respuesta deseable pero poco frecuente en contraposición a la conducta más común pero menos deseable, o bien cuando se elige el reactivo por la conveniencia social de la respuesta.

Dentro de las técnicas o procedimientos más frecuentemente empleados para evidenciar la contaminación de una medición por DS es utilizar una escala explícita de DS en conjunto con la escala de interés y verificar si hay asociación significativa entre ellas; de ser así, se asume la presencia de DS. A este respecto es que se involucra el tema de DS en la evaluación de la capacitación. La valoración del adiestramiento no es nada sencillo, ya que se requiere determinar los cambios en las habilidades, conocimientos y actitudes de los empleados que deben tener lugar cuando se ven sometidos a un proceso llamado “adiestramiento”. Es frecuente que la evaluación de la capacitación sea practicada con

procedimientos desordenados, poco rigurosos y con recursos inadecuados, dándose lugar tan sólo a la recolección de opiniones externas, que carecen de validez y confiabilidad para informar sobre el nivel de madurez del sistema de capacitación, limitando así su desarrollo y consolidación. Para resolverlo, deben crearse medidas confiables y válidas a fin de determinar si el entrenamiento hace en verdad lo que se espera que haga.

Actualmente no hay suficiente evidencias en la literatura científica del efecto de la DS en la distorsión de las respuestas de quienes contestan a los cuestionarios de evaluación de la capacitación, aún cuando se plantea que las decisiones derivadas de éstas, tienen importantes implicaciones para las personas que contestan a dichos cuestionarios (vr. gr. promociones a puestos de ascenso, asignación de nuevas responsabilidades o cambio de funciones y establecimiento de nuevos programas de capacitación). Además, las evaluaciones de capacitación pueden estar afectadas por DS debido a que, aunque se evalúa a un tercero, siempre existe la percepción que se está evaluando al propio empleado, por lo que en el presente estudio tuvo como objetivo probar si un instrumento de capacitación presentaba asociación significativa con una medida criterio de DS y si esto era indicativo de que, incluso cuando se pide la evaluación de una persona (en este caso un instructor), la expresión de la propia opinión también asume un involucramiento de la imagen personal ante una situación de demanda social (ambiente laboral). Así el probar o no que la DS afecta los Métodos de la Evaluación de la Capacitación, se estará en la posibilidad de colaborar en la validez de las medidas utilizadas para la evaluación de programas de capacitación, lo que impacta directamente en el costo-beneficio de la formación y desarrollo del recurso humano en las organizaciones.

En resumen, el propósito del presente estudio fue probar la relación entre las instrucciones socialmente deseables y las medidas de evaluación del desempeño de los instructores en los cursos de capacitación, ya que actualmente no hay evidencias de la distorsión de las respuestas de quienes contestan a los cuestionarios de evaluación de la capacitación, aún cuando se plantea que las decisiones derivadas de éstas, tienen importantes implicaciones para las personas que contestan a dichos cuestionarios (vr. gr. promociones a puestos de ascenso, asignación de nuevas responsabilidades o cambio de funciones y establecer nuevos programas de capacitación en el futuro).

CAPITULO I

EVALUACIÓN DE LA CAPACITACIÓN Y SU BENEFICIOS

La capacitación está orientada a satisfacer las necesidades que las organizaciones tienen de incorporar conocimientos, habilidades y actitudes en sus miembros, como parte de su natural proceso de cambio, crecimiento y adaptación a nuevas circunstancias internas y externas. Compone uno de los campos más dinámicos de lo que en términos generales se ha llamado educación no formal. Sin embargo, la medición de los resultados de las acciones de capacitación en el marco de las organizaciones es, probablemente, el aspecto menos desarrollado de los que componen esta área. Lo anterior se debe a lo difícil, pero posible, que es medir los resultados de la capacitación, la dificultad de operacionalizar y definir claramente las habilidades o conocimientos a medir, por el alto impacto de variables sobre el comportamiento humano y por la magnitud del esfuerzo que significa una medición en materia de tiempo y recursos (Blake, 2000).

Cuando se habla de capacitación se hace referencia a un proceso mediante el cual se busca obtener beneficios determinados que justifican su existencia, por ejemplo: menos accidentes de trabajo, menor rotación, incremento en las aptitudes y mayor productividad o eficiencia organizacional . La capacitación, además de evitar el rezago de las personas, también logra la actualización de las empresas y por ende mayor productividad. Es por eso que cada vez aumenta el número de empresarios preocupados en la capacitación de sus empleados.

Actualmente, muchas organizaciones intentan definir exactamente qué consecuencias presenta la capacitación en su personal. Ya no se conforman con enviar al

personal a un curso y esperar que todo se lleve a cabo como se planea. Las organizaciones desean resultados favorables y desean saber qué habilidades, conocimientos o actitudes han asimilado los empleados en la sesión formativa y hasta qué punto ha cambiado su conducta en los equipos de trabajo y en qué medida ha afectado la formación a los objetivos organizacionales, a fin de contar con la información necesaria para planear los siguientes programas de capacitación y verificar nuevamente sus resultados (Zepeda, 1999).

Uno de los elementos trascendentales que se contempla en la capacitación, es evaluar en forma integral el sistema de capacitación implantado. Su importancia radica en medir el impacto del curso en el escenario laboral, lo que da como resultado el costo beneficio de un curso, en función de la manera que los capacitados aplican los conocimientos obtenidos. Existen varias razones importantes para evaluar la capacitación, como por ejemplo: incluir mejoras en capacitaciones futuras, decidir si la capacitación se imparte al personal más idóneo, si contiene los temas necesarios, etcétera (Grados, 2001).

Sin evaluación, es muy difícil demostrar que la capacitación fue la razón de las mejoras. De la formación de los recursos humanos se puede garantizar la productividad, así como la preparación de los empleados actuales en futuros puestos de trabajo. La evaluación implica la identificación de los correspondientes resultados a medir. Los resultados utilizados en la evaluación de programas de formación incluyen la satisfacción en los programa de capacitación, el aprendizaje de conocimientos o habilidades, el uso de conocimientos teóricos y prácticos en el trabajo, y en los resultados como en las ventas, la productividad, o la prevención de accidentes. La evaluación también puede comparar los costos de capacitación respecto a los beneficios recibidos (retorno de la inversión). Los

resultados utilizados en la formación evaluación ayudarán a determinar el grado en que el programa dio lugar a aprender, y la transferencia de la formación, lo que proporciona información necesaria para mejorarla. Investigaciones previas sobre el impacto de los programas de la capacitación en los Recursos Humanos, han mostrado los beneficios tangibles e intangibles de la capacitación y el retorno de la inversión (véase Bryant. & Terborg., 2008; Fraser, 2007).

La capacitación debe considerarse, como una inversión con un rendimiento del gasto efectuado y con una ganancia para la empresa, por ello debe ajustarse a un esquema eficaz. Por ello es necesario evaluar la capacitación, lo que implica la recolección sistemática de información descriptiva y de juicio que puede usarse para tomar decisiones efectivas en el proceso de capacitación. Tales decisiones incluyen la selección, adopción, modificación y la evaluación financiera de diversas actividades de capacitación.

Los principales propósitos de la evaluación de acuerdo a Landy y Conte (2005) son:

- Determinar si los capacitandos han alcanzado los objetivos del programa
- Proporcionar retroalimentación que pueda mejorar los programas de capacitación para los futuros participantes, mejorando su desempeño y productividad
- Justificar los costos de la capacitación
- Demostrar el valor del proceso de capacitación a los altos directivos indicando si la consecución de los objetivos clave del negocio mejoró después de la capacitación.

El objetivo de la evaluación de la capacitación no es simplemente etiquetar un programa como bueno o malo. Se trata de diseñar un sistema de evaluación sistemático y

riguroso que se aplique a cada una de las etapas del proceso de capacitación, tales como el Pre-test, la Interfase y el Post-test, con el objeto de investigar actitudes, aprendizaje, comportamientos o bien el nivel de conocimiento o información que tiene los capacitados (Bentley, 1993; Grados, 2001). Uno de los modelos más aceptados es el de Kirkpatrick (2000), quien divide la evaluación en cuatro niveles: de reacción, de aprendizaje, de conducta y de resultados. Cabe destacar que la evaluación de las reacciones, si es que se evaluaron apropiadamente, reflejan el sentir y el pensar de los participantes acerca del programa de capacitación. Esta aproximación subjetiva de los que se vivió en el curso es una base perfecta para comprender mejor el aprendizaje y su aplicación en el trabajo. La fase de seguimiento es aquella en la que se evalúa en forma integral el sistema de capacitación implantado y se puede evaluar el impacto del curso en el escenario laboral.

Influyen diversos factores para que la capacitación resulte efectiva, la selección de los empleados, los procedimientos de situación de empleados en sus puestos, las diferencias individuales en aptitudes, temperamento y rapidez de aprendizaje (McGehee & Thayer, 1986), el interés del trabajador en la capacitación, una buena planeación y detección de necesidades de capacitación, entre otros, para que pueda decirse que un curso es adecuado.

Es posible que el efecto del adiestramiento sea negativo, por desgracia, hay muy pocas bases experimentales acerca de los méritos relativos de diferentes procedimientos de evaluación. Esta falta de interés radica en la percepción de la dirección de que evaluación sería un gasto innecesario para comprobar algo que ya asume como eficaz. Para realizar un proceso de evaluación de la capacitación al nivel que sea, se requiere de un marco metodológico experimental u operacional que permita orientar o fundamentar las acciones a

realizar. A través de la metodología experimental se puede establecer un tipo de relación causa – efecto (capacitación-resultados), considerando un grupo experimental, al que se le aplique la capacitación, y un grupo control, al que no se le aplica dicha variable. Esta orientación permite determinar en qué medida la capacitación mejoró, el desempeño, los resultados financieros, la gestión u otro indicador definido previamente, y de esta forma se aísla en gran medida el efecto de otras variables en los resultados, y que son ajenas a la capacitación. La segunda orientación, la operacional, considera aspectos propios de la gestión organizacional en que se determinan los niveles de logro según los diferentes aspectos considerados en él.

Puccio, Firestien, Coyle, y Masucci (2006), reportaron el impacto de la capacitación en la Solución Creativa de Problemas (SCP) dentro de las organizaciones. Su propósito fue el examinar de una forma explícita el modelo de la creatividad aplicada y a su vez demostrar los efectos positivos de la capacitación en la Solución Creativa de Problemas. En su estudio llevado a cabo con estudiantes universitarios, los sujetos fueron evaluados antes, durante y después de la capacitación con pruebas sobre la capacidad cognitivas, personalidad, resolución de problemas y la toma de decisiones técnicas y se encontró que aquellos estudiantes que participaron en los cursos de creatividad mostraron niveles significativamente más altos en el rendimiento a través de un gran número de medidas sobre comportamientos de comunicación en grupos (crítica verbal, apoyo verbal, risa, sonrisa) que quienes no habían tomado la capacitación.

Por su parte Guzmán y Sánchez Escobedo (2006), realizaron un estudio con el propósito obtener evidencia acerca del efecto de un programa de capacitación para

profesores de una institución de educación superior del sureste de México. El grupo experimental se conformó con 219 estudiantes de nueve profesores. Se invitaron a otros nueve grupos paralelos a participar como grupo de control (n = 218). Al principio y al fin del semestre los participantes contestaron la prueba de Pensamiento Crítico California (CCTST por sus siglas en inglés). Los resultados indicaron que aquellos que habían profesores que tomaron la capacitación mejoraron sus estrategias didácticas que aquellos que no tomaron la capacitación

La mayoría de los estudios respecto a los programas de capacitación concentran su atención en la importancia de la evaluación en forma sistemática y cómo los resultados de la misma han ayudado a los directores, los empleados y a los profesionales de los recursos humanos a tomar decisiones sobre la mejora de los programas y los métodos en la capacitación, de los costo-beneficio y del costo-eficacia, elementos de mucha utilidad para el análisis de la rentabilidad de la capacitación. Dewine (1987) reporta el desarrollo y validación del diseño de un instrumento que evalúa las reacciones de los participantes en un programa de capacitación; la aplicación de una aptitud específica en las tareas en trabajo y asimismo evalúa el grado en que los participantes perciben cómo han dominado las habilidades adquiridas.

Según Pinto (2000), es frecuente que la evaluación de la capacitación sea practicada con procedimientos desordenados, poco rigurosos y con recursos inadecuados, dándose lugar tan sólo a la recolección de opiniones externas, que carecen de validez y confiabilidad para informar sobre el nivel de madurez del sistema de capacitación, limitando así su desarrollo y consolidación. Ante esto, Siliceo (2001), argumenta que es

necesario saber aplicar y ejecutar cuatro características y criterios para realizar una buena evaluación, que son:

- Validez. La evaluación se basa en hechos y datos ciertos.
- Confiabilidad. La evaluación tiene alto grado de consistencia al medir y remedir los mismos examinados.
- Objetividad. La prueba está diseñada y construida de manera que pueda ser calificada por dos o más calificadores independientes, dando los mismos resultados.
- Estandarización. La prueba se ha hecho a un grupo de personas suficientemente representativas, en condiciones uniformes, para establecer normas de desarrollo y distribuciones porcentuales de los examinados, con el propósito de hacer comparaciones correctas.

Rodríguez (1993) menciona que los instrumentos de evaluación serán diversos de acuerdo con la naturaleza de los eventos. Éstos pueden contener una serie de reactivos que llevan al participante a emplear la información, las habilidades y las actitudes aprendidas para la solución de determinados problemas. Por esto, un cuestionario que evalué la eficacia de la capacitación estará bien realizado si sus reactivos corresponden a los objetivos y constituyen una buena muestra de los contenidos manejados durante el curso. La mayoría de los instrumentos en cuestión incluyen: a) instrucciones; explicación de lo que tiene que hacer el sujeto y del modo en que debe anotar sus respuestas; b) los reactivos propiamente dichos: las actividades mismas, contestar preguntas, resolver problemas, operar máquinas, etc.

La evaluación de la capacitación es un proceso que debe concebirse en términos de medición válida y confiable. Ha de centrarse en recabar información de varios escenarios laborales en los que ya se hayan implantado programas de capacitación y en donde los

objetivos inicialmente establecidos ya se hayan convertido en procesos realizados habiéndose aplicado por los empleados capacitados, dentro de su contexto de trabajo (Grados, 1998).

No obstante ésta última visión de la capacitación no ha sido la más común y es mucho más generalizada la evaluación a partir de los autoreportes.

CAPITULO II

INSTRUMENTOS DE EVALUACION Y DESEABILIDAD SOCIAL

En la mayoría de las evaluaciones de capacitación el instrumento de uso común es el formulario de evaluación del lápiz-y-papel. La escala tipo Likert ha sido el instrumento más utilizado por los investigadores en los estudios de campo y este tipo, por ser un método eficiente para capturar una amplia gama de respuestas que van desde las actitudes hasta las opiniones. Actualmente existen diversos métodos que evalúan si este tipo de instrumentos, cuentan con las propiedades psicométricas pertinentes para poder realizar generalizaciones adecuadas sobre aquello que se estudia (Benzlein & Berg, 2003; Fineman, 1975; Hartley & MacLean Jr., 2006; Larsen, 1994; Larsen, Reed, & Hoffman, 1980; Meerkerk, Van Den Eijnden, Vermulst, & Garretsen, 2009; Ommundsen & Larsen, 1997; Russell, Lankford, & Grinnell Jr., 1981; Templer, King, Brooner, & Corgiat, 1984). En el campo que nos compete sin embargo, se ha mencionado que la utilización de estos instrumentos de medida arrojan resultados altamente positivos en términos de las calificaciones hacia los instructores. Algunos argumentos que implican que dichos hallazgos deberían ser cuestionados, se debe a que se ha llegado a comprobar los efectos de halo (Darby, 2007; H. C. Smith & Wakeley, 1988), Pigmalión o profecías autocumplidas (Reavis, 1979); la tendencia al acuerdo, la aquiescencia, el error de indulgencia (Chang & Ferre, 1988) y la deseabilidad social (Birenbaum & Montag, 1989; Ferrando, Lorenzo-Seva, & Chico, 2009).

En otros estudios se ha observado por ejemplo que los profesores reciben calificaciones altas debido a la impresión positiva predominante hecha mientras que trabajaban recíprocamente con los estudiantes (Phelps, Schmitz, & Boatright, 1986), que se

reciben pocos comentarios negativos (Chang & Ferre, 1988), debido a que los alumnos esperan recibir las mismas calificaciones que ellos otorgan al profesor o instructor (Ellis, Burke, Lomire, & McCormack, 2003), todo ello debido a la indulgencia.

En el caso particular de las evaluaciones dentro del ambiente laboral, es pertinente aclarar la noción de un contrato psicológico, debido a que posiblemente también sea una fuente de sesgo en las evaluaciones. El contrato psicológico implica la existencia de un conjunto de expectativas, no escritas en parte alguna, que opera a toda hora entre cualquier los miembros y dirigentes de una organización (Schein, 1982). Las relaciones de empleo no se concreta en términos especificados por un contrato laboral de tipo escrito, sino que se inserta en un amplio marco de interacción social, que no incluye solo al trabajador y a su jefe inmediato sino que se extiende más allá, para incluir a los compañeros de trabajo (Palací, 2005), lo que lleva a establecer lazos de camaradería y compañerismo, ya que el simple hecho de permanecer a la misma empresa, los contactos personales o de grupo pueden tener efectos más allá del trabajo (Rüssel, 1963). El contrato psicológico define las condiciones del comportamiento psicológico del empleado con el sistema. Los empleados aceptan conceder una cantidad determinada de trabajo y lealtad (Davis & Newstrom, 1991). Dentro de las expectativas del trabajador está el ser capacitado para el mejor desarrollo de sus funciones. De esta manera el contrato psicológico es una relación de intercambio entre el empleado individual y la organización es una relación implícita basada en contribuciones mutuas (Muchinsky, 2007).

Como se ha mencionado, en el ambiente laboral como en muchos otros escenarios, los evaluadores han de confiar en los reportes como las técnicas de recolección de

información, sin embargo, estas estimaciones pueden distorsionarse por muchos fenómenos dentro de los que se podría también mencionar el manejo de la imagen que se explorará a continuación.

Dentro de la vasta literatura científica, se ha documentado que las diversas herramientas de medición psicológica, particularmente los autoreportes, son susceptibles de contaminarse por diversas variables contingentes denominadas estilos de respuesta. Dentro de estos estilos de respuesta, existe uno en particular denominado Deseabilidad Social (DS) que se refiere a cuando las personas tienden a mostrarse a sí mismos de acuerdo a las normas sociales establecidas de una manera exagerada o poco habitual. La DS implica sobre dimensionar las actitudes y comportamientos socialmente deseables y reducir los comportamientos y actitudes socialmente desaprobados o menos deseables (Collazo, 2005; Coq & Asián, 2002; González & Iruarrizaga, 2005). Las denominaciones frecuentes para este sesgo son otras, tales como “Sinceridad” y “Distorsión Motivacional”. También se han utilizado denominaciones como “Simulación”, “Virtudes Infrecuentes” o “Distorsión de la Respuesta”. De todas estas denominaciones, “DS” es la más frecuentemente utilizada y todas las demás caben dentro de esta etiqueta (Salgado, 2005, p. 118).

Las personas se inclinan a dar una respuesta de sí mismos de tal manera que se vea bien en una auto-descripción o al menos en una situación en donde su imagen se ve implicada, lo que origina que los reactivos de una prueba de personalidad se contesten conforme a las presiones sociales o normativas en lugar de proporcionar un autoinforme verídico (Li & Bagger, 2007). En otras palabras, es la tendencia a escoger las respuesta deseable reconocida comúnmente, que está determinada por la conveniencia social de la

respuesta (Braun, Jackson, & Wiley, 2002; Kerlinger & Howard, 2002; Nunnally & Bernstein, 1995).

El interés acerca de la DS se inicia en los años 30 con los trabajos de Bernreuter (1933), alcanza su máximo apogeo entre los años 1955 y 1965 y parece que vuelve a reavivarse desde principios de los 90. El punto de partida del constructo DS es el supuesto de que algunas de las alternativas de respuesta en ciertos reactivos de personalidad son socialmente más convenientes o deseables que otras, por lo que algunos sujetos pueden tender a elegir las independientemente de cuál sea su nivel en el rasgo que el reactivo pretende medir (Ferrando & Chico, 2000). Más en detalle, la teoría inicial planteaba que ciertos sujetos podían dar, voluntariamente, una imagen distorsionada de sí mismos respondiendo siempre a las alternativas socialmente más deseables de los reactivos. Lógicamente, se planteaba también que esta tendencia a ‘disimular’ o a ‘quedar bien’ se acentuaría cuando la motivación para hacerlo fuese alta, por ejemplo en situación de selección para un puesto de trabajo. La mentira, el engaño, la ocultación, la deformación y tergiversación de la información son conductas comunes al ser humano, que persiguen diversos fines en el proceso de interacción social como ofrecer una impresión positiva a los demás, auto-beneficiarse y beneficiar a otros, evitar un castigo potencial o simplemente mantener una buena interacción social procurando no herir innecesariamente los sentimientos de los demás (Martínez Selva, 2005). Por su parte Homanns, (1974 en Harry, 1984) ha sugerido que la DS tiene su origen en la necesidad de la aprobación social, la cual es concebida como un factor de motivación importante en el comportamiento humano y que hay una disposición general de parte de individuos para buscar la evaluación favorable de otros.

Anastasi y Urbina (1998), indican que hay personas que se sienten motivadas a distorsionar sus respuestas para “quedar bien” ante los demás, es decir, elegir respuestas que generen una impresión favorable, como al solicitar empleo o al buscar la admisión en una institución educativa, al igual que hay personas que motivadas a distorsionar sus respuestas para “quedar mal”, tratan de aparentar un trastorno psicológico mayor al que presentan, por ejemplo en la evaluación de quienes están sometidos a un juicio por haber cometido un crimen. En ambos casos, la distorsión implica que las mediciones que se hacen no son confiables y pueden llevar al evaluador a conclusiones erróneas o diagnósticos mal definidos, de tal suerte que se hace necesario poder identificar a aquellos sujetos que de manera intencional distorsionan sus respuestas.

Operacionalmente, se considera que una persona esta contestando con DS, cuando su tasa de respuestas en un test se aparta significativamente del patrón de respuestas esperado o normalizado en una dirección favorable, asumiendo que la persona está dando una imagen de sí misma distorsionada o falsa. En estos casos, el test se considera inválido y esto impide poder hacer algún tipo de inferencia sobre la conducta real de la persona (Cabañeros, García Cueto, & Lozano, 2003).

Principalmente la DS es una fuente de invalidez cuando directamente se involucra la auto-imagen de la persona evaluada; es decir, cuando se requiere que las personas digan algo acerca de ellas mismas. Se ha asumido que una técnica más general para evitar este efecto es parafrasear los reactivos en tercera persona y esto hace que disminuya el involucramiento, haciéndose la situación menos amenazante para el respondiente y por ende, proporciona más probabilidades de que conteste sin DS, aunque lo más conveniente

es utilizar el anonimato cuando es posible (Paulhus & Reid, 1991; Stöber, Dette, & Musch, 2002; Van de Vijver, 1997, 1998), y pedirles completa honestidad Gordon (1987).

No obstante, cuando el sujeto se encuentra en una situación de demanda social y no es posible el anonimato, la DS se presenta, incluso aunque la evaluación sea de terceros, ya que al hablar mal o bien acerca de otros también damos una impresión de nosotros hacia los demás (M. E. Gordon & Petty, 1971; Konstabel, Aavik, & Allik, 2006).

Desde luego existen puntos de vista opuestos en el que se sustenta que no hay claridad sobre la frecuencia de la distorsión. Landy y Conte (2005) discuten que existen reportes con una tasa de falsedad importante y en otros en donde indican que es mínima y ofrece un ejemplo de un estudio donde los participantes recibieron instrucciones para distorsionar sus respuestas con el fin de hacerlos ver más favorables y al comparar los resultados, la validez predictiva de las medidas de personalidad permaneció sin cambio.

La DS consiste en dos dimensiones, que han sido denominadas manejo de impresión (impression management) y auto-engaño (self-deception) (Paulhus, 1991). El manejo de impresión indica una tendencia a adaptar intencionalmente la imagen pública de uno mismo con objeto de obtener una visión favorable por parte de los demás. El auto-engaño, por el contrario, se refiere a la tendencia no intencionada de describirse a uno mismo de un modo favorable y que se manifiesta en auto-descripciones sesgada positivamente pero en las que se cree honestamente. En este sentido, el manejo de impresión es una manipulación voluntaria de la propia imagen para que los demás lo perciban a uno de modo positivo, mientras que el auto-engaño no es una manipulación

deliberada aunque pueda conducir a distorsiones en la percepción que los otros tienen de uno mismo (Salgado, 2005).

Kuncel y Tellegen (2009), definen a la respuesta socialmente deseable, como un comportamiento, que es percibido como lo más compatible y deseado por otros cambiado de situación a la situación (el trabajo, la escuela, el matrimonio), lo que ocasiona que haya comportamientos muy deseables en algunas situaciones mientras que en otras, la misma conducta es menos deseable. De igual forma, Kuncel y Tellegen hacen alusión a que algunos comportamientos están diseñados para dar una impresión favorable bajo una rutina ya hecha y que existen comportamientos automáticos e inconscientes, pero que otros pueden ser deliberadamente manipulados y resultar engañosos. En el uno o el otro caso la conducta está diseñada para efectuar algún grado de engaño, con el objetivo de influir en las decisiones o las opiniones de otras personas.

Estas distorsiones intencionadas, también pueden llevarse a cabo en una dirección poco deseable. Hay evidencia que indica que algunos individuos pueden sentirse motivados a “falsear para quedar mal”, tratando de aparentar un trastorno psicológico mayor al que presentan, por ejemplo en la evaluación de quienes están sometidos a un juicio por haber cometido un crimen. Rogers et al (2002), revelaron un efecto contrario, en un estudio con delincuentes adolescentes. Deseaban comprobar si podrían intencionalmente modificar sus grados de la psicopatía para alcanzar un propósito deseado, adoptando cualquiera papel pro-social con respeto a la autoridad (es decir, DS) o a un papel de criminal-resistente que desafía a la autoridad (es decir, inconformidad social). Los resultados demostraron cómo

los delincuentes reducirían al mínimo su psicopatía y se presentaban con un estilo social deseable de la respuesta.

Bartlett y Richard (1967), examinaron diversos estudios en los cuales, indican que el procedimiento más común para probar la susceptibilidad a la falsificación en medidas de personalidad, es administrar el cuestionario bajo cierta clase de instrucciones a fin de falsear las respuestas. Basado en ello llevaron a cabo un estudio cuyo propósito fue determinar las diferencias entre los estilos de la respuesta que son obtenidos por diferentes instrucciones. Eligieron tres condiciones para el estudio, siendo éstas: (1) instrucciones para contestar honestamente con poca motivación a falsear, (2) instrucciones para contestar honestamente con considerable motivación a la falsificación y (3) instrucciones específicas para falsificar. Los resultados demostraron que las instrucciones para contestar de una manera más positiva (instrucciones para falsificar), condujeron a responder considerablemente más alto en DS que cualquiera de las instrucciones para contestar honestamente.

Al igual que con cualquier medida de lápiz y papel de las características y comportamientos humanos, un problema relacionados con los inventarios de personalidad tiene que ver con su validez. Muchos de los reactivos de los inventarios requieren que quienes responden admitan cosas que podrían desear no admitir, sino más bien presentarse de la manera más favorable. La veracidad al responder puede ser un problema grave y es probable que los individuos no estén dispuestos a decir la verdad (Aiken, 2003).

Muchos investigadores han buscado la forma de medirla y, si es necesario, controlarla para asociarla a las distorsiones en los auto-reportes. Una importante alternativa

en esta tarea, ha sido el desarrollo de escalas para medir las diferencias individuales en la respuesta socialmente deseable. A través de los años, numerosas de estas escalas han sido desarrolladas y aplicadas ampliamente en diversos campos de investigación básica; sin embargo, sus usos en la práctica, han sido cuestionados por la forma cómo se debe evaluar y controlar la DS en los inventarios auto-descriptivos.

A continuación se mencionan las dos escalas más utilizadas dentro de la literatura que se han desarrollado y probado para medir la DS: Escala Marlowe y Crowne y el Inventario Balanceado de Conductas Deseables (actualmente conocido como BIDR)

La Escala de Deseabilidad Social de Marlowe-Crowne (M-C Social Desirability, MCSDS)

Hace 40 años, Crowne y Marlowe (1960) desarrollaron una escala para medir la DS modificando alguno de los supuestos propuestos inicialmente por Edwards (1957) quien fuera el primero en acuñar el término de DS. Los autores definieron a la DS como aquellas conductas aprobadas culturalmente pero con poca ocurrencia y las conductas desaprobadas pero frecuentes. Basándose en esta definición, seleccionaron ítems de algunos cuestionarios de personalidad vigentes en esa época que cumplieran con la definición propuesta. Entonces, una muestra de estudiantes universitarios clasificó el grado de DS de cada uno de los ítems seleccionados. Posteriormente se conservaron solamente los ítems que discriminaban entre los marcadores más altos y bajos en un nivel de .05, y juzgados ulteriormente para tener menos desequilibrio relativo al contenido de la escala de DS de Edwards. Este conjunto de los ítems parecían tener buenas características psicométricas (consistencia interna = 0.88, test-retest $r = 0.89$), por lo que los autores llegaron a una

nueva escala, la escala de DS de Marlowe y Crowne, conocida hoy en día como MCSDS por sus siglas en inglés (O'Grady, 1988).

La MCSDS es una de las distintas escalas que se aplican para evaluar diferencias individuales en DS. La versión de la MCSDS que se conoce actualmente está constituida por 33 ítems que se contestan por elección forzada entre Verdadero (V) y Falso (F). A mayor puntaje corresponde mayor DS. 18 de los ítems son denominados como atribución. Los ítems de atribución se refieren a la aprobación social pero de conductas no comunes. El resto de los ítems son considerados de negación y se refieren a conductas socialmente impropias pero comunes. En los ítems 1, 2, 4, 7, 8, 13, 16, 17, 18, 20, 21, 24, 25, 26, 27, 29, 31 y 33 a cada respuesta V le corresponde el puntaje 1 y a cada respuesta F le corresponde el puntaje 0. En los ítems restantes se puntúa de forma inversa. El puntaje total se obtiene con la suma de las respuestas de cada ítem. La MCSDS ha sido utilizada para evaluar los sesgos de respuestas en un sentido socialmente deseable en diversas áreas de investigación tanto en Medicina, Mercadotécnica y Psicología (Beretvas, Meyer, & Leite, 2002; Ramanaiah & Martin, 1980).

Johnson y Fendrich (2002) han encontrado estudios en dónde el MCSDS se correlaciona perceptiblemente con una variedad de medidas de psiquiátricas y psicológicas. Por ejemplo, señalan ciertas investigaciones que demuestran constantemente una asociación negativa entre los síntomas de una salud mental pobre y las medidas del MCSDS, lo que ha sido interpretado como una clara evidencia de cómo personas con cualidades indeseables, reportan una alta necesidad de aprobación. También se ha encontrado que se correlaciona inversamente con reportes en el uso de sustancia y correlaciones positivas significativas

entre la escala del MCSDS y las autoevaluaciones favorables, incluyendo la felicidad, la satisfacción de la vida y la moral.

A pesar de que el MCSDS ha sido ampliamente usado por ser una medida fácil y relativamente sencilla de utilizar para estimar DS, también se han realizado diversas críticas a ella. Stöber (2001), argumenta que algunos contenidos de la escala del Marlowe-Crowne pueden ser anticuados. Se cree que los ítems como: “ha habido ocasiones en las que he tenido ganas de revelarme en contra de personas con autoridad aun cuando sabía que tenían la razón” y “en ocasiones he insistido arduamente para que las cosas sean como yo quiero”, pudieran reflejar estándares sociales de finales de la década de 1950, justo cuando la escala de Marlowe-Crowne fue construida y no reflejar necesariamente los estándares actuales.

Lo cierto es que el MCSDS es una de las escalas para medir y controlar el sesgo en los estilos de respuestas en los auto-reportes, desde su aparición se han propuesto y adoptado diversas versiones cortas en el campo de la psicología y la medicina, en una gran variedad de poblaciones sin la pérdida significativa de confiabilidad (Avery & Gary, 1985; Barger, 2002; Loo & Thorpe, 2000; Ray, 1984; Reynolds, 1982; Seol, 2007; Strahan & Gerbasi, 1972; Zook II & Sipps, 1985).

The Balanced Inventory of Desirable Responding (BIDR)

El Inventario Equilibrado de Respuesta Deseable, BIDR por sus siglas en inglés, es un descendiente del Cuestionarios del Autoengaño y Engaño a otros (Self- and Other-Deception) desarrollado por Sackeim y Gur en 1978. Los ítems originales de autoengaño fueron desarrollados suponiendo que aquellos individuos con una propensión para el autoengaño, niegan tener pensamientos o sentimientos psicológicamente amenazantes.

Dichas amenazas estaban basadas en la teoría psicoanalítica (vr. gr. odiar a los padres, disfrutar de evacuaciones intestinales o tener fantasías sexuales). El enfoque de los ítems ha cambiado de la defensa de ego al realce (a la mejora) del ego. La versión más reciente creada por Paulhus en 1988, enfatiza atributos cognoscitivos de manera positiva, es decir el exceso de confianza en los juicios de uno mismo (Paulhus, 1991).

El BIDR mide dos constructos el cual Paulhus los denominó como autoengaño (AE) y manejo de impresiones (MI). El AE se refiere a una distorsión inconsciente positiva en las respuestas a los reactivos con la finalidad de proteger el auto concepto; es la tendencia de dar auto reportes que son honestos pero posiblemente distorsionados. Kashy y DePaulo (1996), lo definen como la tendencia no intencionada de describirse a uno mismo de un modo favorable y que se manifiesta en auto-descripciones sesgadas positivamente pero en las que se cree honestamente. El término de autoengaño ha sido usado en al menos tres distintas maneras: como debilidad moral para rechazar alguna verdad desagradable; como un estado inherente a los mecanismos de defensa; y como término genérico para cubrir la mala percepción de uno. En amplio sentido el concepto de autoengaño ha sido aplicado en la etología humana así como psicología social, clínica y de personalidad. En la versión más reciente del BIDR, el autoengaño ha sido particionado en el realce y la negación. Las investigaciones que han usado el instrumento han confirmado que el primer componente de la varianza de la DS, encontrado en los auto reportes, es el realce del autoengaño (Paulhus, Fridhandler, & Hayes, 1997).

En contraste, el MI se refiere a una disimulación consciente en la respuesta al reactivo con la finalidad de fabricar una impresión favorable en las personas. El manejo de

impresiones indica una tendencia a adaptar intencionalmente la imagen pública de uno mismo y así obtener una visión favorable por parte de los demás. En este sentido, el manejo de impresiones es una manipulación voluntaria de la propia imagen para que los demás lo perciban a uno de modo positivo mientras que el auto-engaño no es una manipulación deliberada aunque pueda conducir a distorsiones en la percepción que los otros tienen de uno mismo.

Los ítems originales de autoengaño fueron racionalmente desarrollados en la creencia de que los individuos con una propensión para el autoengaño tienden a negar pensamientos o sentimientos amenazadores. Los reactivos del manejo de impresiones están racionalmente desarrollados en la creencia de que algunos respondientes sobre-reportan sistemáticamente su actuación en una amplia variedad de comportamientos deseables y sub-reportan los comportamientos indeseables. Ya que las afirmaciones están envueltas de sobre-comportamientos cualquier distorsión es presumiblemente una mentira consciente. El BIDR cuenta con 20 reactivos de AE y 20 de MI, balanceados en igual proporción en sentido positivo y negativo, de ahí su nombre *inventario balanceado*. El BIDR goza de un amplio uso dentro de la comunidad científica por su popularidad tanto en los campos básicos como en los aplicados de la investigación psicológica (Paulhus, 1984).

Las escalas de DS se utilizan sobre todo para ayudar a proporcionar evidencia de la validez de las respuestas en los exámenes psicológicos. El uso de estas escalas implica el cálculo de correlaciones y análisis factoriales entre los puntajes de la escala de DS y los puntajes del instrumento psicológico. El uso más frecuente de las escalas de DS ha sido para invalidar mediciones, a través de la identificación de altas correlaciones entre las

escalas de interés y las escalas de DS como criterio de validación externo. Así, la toma de decisión se basa en eliminar las respuestas hechas por los participantes con elevadas calificaciones en las medidas de la DS.

En algunos casos, diversos investigadores han evaluado la capacidad predictiva y explicativa de la DS para resolver importantes problemas del comportamiento humano (Armitage, 1999; Lewis, 2000; S. Smith, Smith, & Seymour, 1993). A través de la aplicación del análisis de regresión múltiple, se han propuesto distintos modelos que involucran significativamente a la DS en la explicación de la varianza en los autoreportes. Mills, Loza, & Kroner (2003), mostraron que la administración de la impresión tiene una importante contribución relativa en la predicción de la reincidencia de conductas violentas ($R^2 = 0.60$, $p < 0.001$). Los autores proponen que la administración de la impresión puede ser una característica duradera en una muestra de delincuente, más bien que un estilo circunstancial de respuesta. La DS también ha sido indicada como potencialmente influyente en la expresión de actitudes negativas hacia inmigrantes en Australia, en un 8% de la varianza (Schweitzer, Perkoulidis, Krome, Ludlow, & Ryan, 2005). Existen resultados de autoreportes que indican que las escalas DS de Jackson (1984), MCSDS, (1960) y el género predicen en su conjunto hasta un 25% las medidas de adaptación de estudiantes a la universidad (Silverthorn & Gekoski, 1995).

El poder predictivo que se ha encontrado va de los problemas más generales a los más particulares, relacionando en cada caso un factor o factores con resultados específicos. A pesar del amplio uso que se le ha dado a las escala de la DS para la validación de otros instrumentos, algunos opinan que los resultados de las escalas aún no ha sido lo

suficientemente validadas, argumentando que muchos de los estudios no han utilizado apropiadamente las técnicas estadísticas que prueben su dimensionalidad (Leite & Beretvas, 2005). Sin embargo, a pesar de las críticas y falta de evidencia contundente de su poder predictivo, hasta el momento, son las alternativas más eficaces para inferir la validez de las mediciones a través de autoreportes.

CAPITULO III

MÉTODO

Si los cuestionarios que evalúan el desempeño de los instructores de capacitación, presentan respuestas distorsionadas de modo positivo, se corre el riesgo de evaluar de manera favorable los programas de capacitación, influenciado por el efecto de la DS, lo que conlleva medir deficientemente el grado en que se lograron los objetivos del aprendizaje, repercutiendo significativamente en el costos-beneficio y el costo-eficacia, elementos de que permiten determinar la rentabilidad de la capacitación.

El objetivo del presente estudio fue probar si un instrumento de capacitación presentaba asociación significativa con una medida criterio de DS y si esto era indicativo de que, incluso cuando se pide la evaluación de una tercera persona (en este caso un instructor), la expresión de la propia opinión también asume un involucramiento de la imagen personal ante una situación de demanda social (ambiente laboral).

Planteamiento del Problema

¿Existe relación con el cuestionario que evalúan el desempeño de los instructores de la capacitación (CEC) y las escalas del MCSDS y BIDR, que miden la DS?

¿Qué relación existe entre el CEC y las escalas del MCSDS y BIDR, cuando se agrupan por tipo de instrucción (Evaluación favorable inducida o tipo estándar)?

¿Qué relación existe entre el CEC y las escalas del MCSDS y BIDR, cuando se agrupa por tipo de instrucción (Evaluación favorable inducida o tipo estándar) y por tipo de instructor (interno o externo)?

¿Si los participantes a cursos de capacitación, reciben instrucciones que inducen evaluar favorablemente el desempeño de los instructores internos, sus puntajes de las evaluaciones diferirán de los que siguen instrucciones de tipo estándar?

¿Si los participantes a cursos de capacitación, reciben instrucciones que inducen evaluar favorablemente el desempeño de los instructores externo, sus puntajes de las evaluaciones diferirán de los que siguen instrucciones de tipo estándar?

¿Existe una relación lineal entre los puntajes del CEC y las subescalas del BIDR y MCSDS?

Hipótesis

H₁ El cuestionario que evalúa el desempeño de los instructores de la capacitación (CEC) y las escalas del MCSDS y BIDR, que miden la DS, se relacionan positiva y significativamente.

H₂ El CEC y las escalas del MCSDS y BIDR, cuando se agrupan por tipo de instrucción (Evaluación favorable inducida o tipo estándar), se relacionan positiva y significativamente.

H₃ El CEC y las escalas del MCSDS y BIDR, cuando se agrupan por tipo de instrucción (Evaluación favorable inducida o tipo estándar) y por tipo de instructor (interno o externo), se relacionan positiva y significativamente.

H₄ Los puntajes de las evaluaciones de los participantes a cursos de capacitación, que reciben instrucciones que inducen evaluar favorablemente el desempeño de los

instructores internos y externos, difieren significativamente de los puntajes de quienes siguen instrucciones de tipo estándar.

H₅ Existe una relación lineal en los puntajes del CEC y las subescalas del BIDR y MCSDS.

Diseño

El presente estudio es un experimento de campo transversal-correlacional a fin de determinar la asociación entre los puntajes de la evaluación del desempeño de los instructores de la capacitación, respecto a la DS, al término de cada curso de capacitación.

Participantes

Los participantes de la presente investigación fueron empleados de una empresa paraestatal con sede en ciudad de México, que asistieron a un curso de capacitación, de tipo administrativo, técnico, informático, seguridad física y de desarrollo humano (20 horas mínimo y 40 horas máximo). El tamaño de la muestra fue de 280 participantes, 167 mujeres y 113 hombres, seleccionados de manera no probabilística intencional. Los sujetos fueron asignados aleatoriamente entre los dos métodos de recogida de información utilizados (ver procedimiento). Para el primer método de recogida de datos (Instrucción-Evaluación favorable inducida) participaron 138 sujetos, 80 mujeres con una media de edad de 39.68 años y una desviación típica de 10.39, rango de 21 a 62 años y 58 hombres con una media de edad de 43.55 años y una desviación típica de 10.52, rango de 25 a 66 años (uno no proporcionó sus datos). El segundo grupo (Instrucción-Estándar) estuvo formado por 140 sujetos, 87 mujeres con una media de edad de 43.06 años y una desviación típica de 10.84, rango de 21 a 72 años y 53 hombres con una edad media de 44.75 años y una desviación

típica de 10.25, rango de 18 a 65 años (sólo uno no proporcionó sus datos). La media de edad para la muestra total fue de 42.51 años y una desviación típica de 10.65, rango de 18 a 72 años. En cuanto a la situación contractual la mayoría de los trabajadores eran de base (definitivos); 93.73% sindicalizados y solo un 6.27% de confianza. En el formato del CEC, no se contempla recabar los datos de escolaridad, ni antigüedad del trabajador.

Instrumentos

Para llevar a cabo la investigación se utilizaron los siguientes instrumentos.

El Cuestionario para la Evaluación de la Capacitación (CEC): Es una escala tipo Likert, compuesta de 9 reactivos que evaluación al desempeño del instructor al término de cada curso de capacitación. Los reactivos se presentan en una escala de respuesta de 5 puntos que van desde 6 (Totalmente en desacuerdo) a 10 (Totalmente de acuerdo). El CEC tiene un coeficiente de confiabilidad $\alpha = 0.92$ y una varianza acumulada de 72.61%. Este cuestionario ha sido utilizado desde 30 años y fue desarrollado explícitamente por la institución para el área de recursos humanos.

Escala de Deseabilidad Social de Crowne y Marlowe (1960). Contiene 33 ítems de verdadero y falso. Los primeros dieciocho ítems se refieren a conductas aprobadas pero poco frecuentes y los otros 15 ítems se refieren a conductas indeseables pero frecuentes. Las confiabilidades estimadas mediante índices de consistencia interna oscilan entre 0.75 y 0.85, para éste estudio se obtuvo un índice de 0.76. La estructura latente de la escala mostró una varianza acumulada de 53.65%. La matriz de componentes rotados, muestra una solución de dos factores la cual cuenta con los mismos 18 reactivos considerados de atribución y los 15 reactivos de negación del MCSDS (Ramanaiyah & Martin, 1980).

El Inventario Equilibrado de Respuestas Deseables (Balanced Inventory of Desirable Responding, BIDR, Paulhus, 1994) de 40 ítems, que mide los dos aspectos de la DS: Manejo de la impresión de 20 ítems, que se refiere a la disimulación consciente de la respuesta al ítem, con el objetivo de dar una impresión favorable hacia otros y el autoengaño de 20 ítems, que se refiere a la inconsciente tendencia a responder a ítems de manera positiva, con el objeto de presentarse de manera positiva. Los ítems se presentan con una escala de respuesta de 5 puntos en los límites de 1 (no es cierto) a 5 (muy cierto). El BIDR tiene un coeficiente de confiabilidad que se extiende de 0.68 a 0.80 para la subescala del autoengaño y a partir la 0.75 a 0.86 para la subescala de administración de la impresión. En éste estudio el índice fue de 0.80 para la escala total.

Procedimiento

Cada año la empresa en la que se llevó a cabo el estudio programa una serie de cursos de capacitación para el personal de la gerencia de recursos materiales y servicios generales de las oficinas del centro administrativo de la ciudad de México. Se identificaron los cursos que concluirían en cada una de las semanas en que se llevó a cabo durante febrero a septiembre de 2009. Una vez identificados los cursos dentro del periodo y previo a la clausura de cada evento, se prepararon paquetes, de acuerdo al número de participantes por curso, los cuales contenían el Cuestionario de Evaluación de la Capacitación (CEC), la escala de Marlowe y Crowne (MCSDS) y el Inventario Equilibrado de Respuestas Deseables (BIDR), ordenados aleatoriamente. Al CEC se le incorporó al azar una de las instrucciones, tanto para el instructor interno como al externo, ver anexo A) y se entregó un paquete a cada uno de los participantes también de manera aleatoria. El lugar seleccionado para la aplicación fue el aula en donde se llevó a cabo el evento, con espacio suficiente para

alojar a 20 participantes a la vez. La aplicación de los cuestionarios fue realizado por un prestador de servicio social de la carrera de psicología, instruido previamente para dar las siguientes indicaciones para el llenado de los instrumentos: a) en cada sesión se explicó al personal asistente el objetivo del cuestionario, b) escribir su número de empleado y el nombre del instructor y el nombre del evento, en las casillas correspondiente, y c) se les pidió que leyeran y siguieran las instrucciones, a fin de asegurarse de que quedara clara la forma en la que se debería responder a cada una de las preguntas. Se aseguró que durante la aplicación de los cuestionarios, éstos se realizaran de manera individual y se respondieran a todos los ítems. Una vez contestados los cuestionarios se procedió a capturar los resultados y su análisis.

Tratamiento Estadístico

A fin de determinar si existía una relación entre los puntajes de las evaluaciones del desempeño de los instructores de cursos de capacitación y los puntajes de las escalas del MCSDS y BIDR, se utilizó la prueba de correlación de Spearman y posteriormente se segmentaron, en primera instancia, por tipo de instructor (interno o externo) y por tipo de instrucción (Evaluación favorable inducida o tipo estándar) para observar si existía alguna variación entre éstas. Para comparar los promedios de las evaluaciones del desempeño de los instructores externo e interno se aplicó una t de Student para muestras independientes y adicionalmente se corrió la misma prueba segmentada esta vez por tipo de instrucción (Evaluación favorable inducida o tipo estándar). Para probar los efectos principales por tipo de instructor y tipo de instrucción, se realizaron análisis de varianza de dos vías (ANOVA). Por último y a fin de conocer la capacidad predictiva de las escalas de DS (BIDR y

MCSDS) y sus subescalas, en los puntajes del CEC, se llevó a cabo un análisis de regresión múltiple por el método de pasos sucesivos.

CAPITULO IV

RESULTADOS

La tabla 1 muestra la carga factorial del Cuestionario de Evaluación de la Capacitación en donde se obtuvo 1 factor: (F1) “Evaluación del desempeño del instructor”, que habla de la didáctica del instructor, la integración de equipos de trabajo, conocimientos y habilidades para el logro de los objetivos del curso, con una varianza total de 72.61% ($\alpha = .92$).

Tabla 1

Carga factorial del Cuestionario de Evaluación de la Capacitación para una solución ortogonal varimax de un factor

Reactivo	Carga factorial
Factor 1.- Evaluación del desempeño del instructor	
2 ¿Desarrolló el tema en forma clara y comprensible?	.888
9 ¿Considera que se cumplió el objetivo del evento?	.885
4 ¿Resolvió todas sus dudas en forma clara?	.879
1 ¿En qué grado demostró dominio del tema?	.867
6 ¿Permitió la intervención de los participantes?	.847
3 ¿En qué grado despertó su interés en los temas expuestos?	.842
5 ¿En qué grado relacionó los temas con aplicaciones en el trabajo?	.827
7 ¿Como transmitió el conocimiento en relación al Objetivo del Curso?	.822
8 ¿Inició y terminó puntualmente las sesiones?	.810

A continuación se muestran en la tabla 2 el promedio de calificaciones de cada una de las escalas utilizadas, su desviación estándar, el puntaje de consistencia interna y las intercorrelaciones entre las escalas y el cuestionario de evaluación de la capacitación. Inicialmente se puede observar la correlación esperada entre las escalas de MCSDS y BIDR, ya que esta última incorpora reactivos del Autoengaño ($r = .55$, $p = .01$). Se encontraron correlaciones bajas pero significativas, respecto al CEC y las escalas de DS (BIDR, $r = .360$, $p = .01$; MCSDS, $r = .235$, $p = .01$). Por otra parte la media de la escala de MCSDS (23.24), muestra ser superior respecto a los reportes de la versión original $M = 13.72$ y $SD = 5.78$ y se confirma nuevamente que en los estudios realizados en México se obtienen medias bastante más altas (véase Domínguez, Salas, & Reyes-Lagunes, 2008).

Tabla 2
Intercorrelaciones y estadísticos básicos del MCSDS, BIDR, CEC.

Escala/Cuestionario	Rango	M	DE	α	Correlaciones					
					1	a ¹	b ¹	2	a ²	b ²
1. MCSDS	0 a 33	23.24	4.95	.768	-	-	-	-	-	-
a ¹ . Atribución	0 a 18	12.39	3.07	.672	.799**	-	-	-	-	-
b ¹ . Negación	0 a 15	10.83	3.10	.757	.752**	.250**	-	-	-	-
2. BIDR	0 a 40	17.17	3.49	.804	.556**	.371**	.521**	-	-	-
a ² . Auto engaño	0 a 20	6.06	4.95	.772	.375**	.251**	.362**	.845**	-	-
b ² . Manejo de la impresión	0 a 20	11.11	3.84	.686	.572**	.384**	.527**	.814**	.398**	-
3. CEC	6 a 90	86.91	5.04	.953	.235**	.222**	-	.360**	.304**	.295**

Nota. N = 263 para las Intercorrelaciones de CEC y MCSDS. N = 280 para las Intercorrelaciones de CEC y BIDR. MCSDS = Escala de DS de Marlowe-Crowne, BIDR = Balanced Inventory of Desirable Responding, CEC = Cuestionario de Evaluación de la Capacitación. α = Cronbach's α , ** $p < .01$.

En la tabla 3 se observan las intercorrelaciones entre el CEC y la escalas de MCSDS y BIDR, segmentada por tipo de instrucción, en donde se observa una correlación significativamente más altas con las instrucciones de tipo estándar (MCSDS, $r = .246$, $p < .01$; BIDR, $r = .389$, $p < .01$). Mientras que con instrucciones que inducen evaluar favorablemente el curso, el CEC mostró una relación moderada pero significativa con el MCSDS de $r = .223$, $p < .05$ y con el BIDR de $r = .335$, $p < .01$

Tabla 3
Intercorrelaciones del CEC, MCSDS y BIDR en función al tipo de instrucción.

Instrucción	Correlaciones											
	Evaluación favorable inducida ¹						Estándar ²					
Escala/Cuestionario	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1. MCSDS	-	-	-	-	-	-	-	-	-	-	-	-
a ¹ . Atribución	.826**	-	-	-	-	-	.775**	-	-	-	-	-
b ¹ . Negación	.774**	.320**	-	-	-	-	.731**	.191*	-	-	-	-
2. BIDR	.617**	.405**	.627**	-	-	-	.501**	.341**	.421**	-	-	-
a ² . Auto engaño	.451**	.295**	.470**	.833**	-	-	.329**	.231**	.272**	.861**	-	-
b ² . Manejo de la impresión	.594**	.409**	.586**	.813**	.377**	-	.541**	.348**	.469**	.818**	.434**	-
3. CEC	.223*	-	.239**	.335**	.291**	.258**	.246**	.304**	-	.389**	.323**	.341**

Nota. CEC = Cuestionario de Evaluación de la Capacitación. MCSDS = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación. BIDR = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. Instrucción Evaluación favorable inducida¹ N = 123; Evaluación Estándar² N = 140. * $p < .05$, ** $p < .01$.

En la tabla 4 se observa que con los instructores internos, el CEC se correlaciona significativamente con el BIDR, tanto con las instrucciones que inducen evaluar favorablemente ($r = .484$, $p < .01$) como con las estándar ($r = .370$ $p < .01$). El CEC y el MCSDS mostró una relación baja pero significativa con las instrucciones estándar ($r = .296$ $p < .01$) y una correlación media con la subescala de atribución de .424.

Tabla 4
Interrelaciones del CEC, MCSDS y BIDR en función al tipo de instrucción e instructor interno

Instrucción	Evaluación favorable inducida						Estándar					
	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1. MCSDS	-	-	-	-	-	-	-	-	-	-	-	-
a ¹ . Atribución	.779**	-	-	-	-	-	.777**	-	-	-	-	-
b ¹ . Negación	.715**	-	-	-	-	-	.707**	-	-	-	-	-
2. BIDR	.569**	.354**	.568**	-	-	-	.482**	.324**	.394**	-	-	-
a ² . Auto engaño	.440**	.270*	.467**	.858**	-	-	.359**	.250*	.287**	.863**	-	-
b ² . Manejo de la impresión	.544**	.363**	.507**	.782**	.384**	-	.512**	.332**	.426**	.838**	.478**	-
3. CEC	-	-	.291*	.484**	.428**	.365**	.296**	.424**	-	.370**	.291**	.339**

Nota. CEC = Cuestionario de Evaluación de la Capacitación. MCSDS = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación. BIDR = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. Correlaciones por instructor interno vs Evaluación favorable inducida N = 67; Correlaciones por instructor interno vs Instrucción Estándar N = 94; * $p < .05$, ** $p < .01$.

Ambas escalas no mostraron relación con el CEC al segmentarse por instructor externo y por las instrucciones inducidas y estándar (tabla 5).

Tabla 5												
Intercorrelaciones del CEC, MCSDS y BIDR en función al tipo de instrucción e instructor externo												
Instrucción	Evaluación favorable inducida						Estándar					
	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1. MCSDS	-	-	-	-	-	-	-	-	-	-	-	-
a ¹ . Atribución	.870**	-	-	-	-	-	.740**	-	-	-	-	-
b ¹ . Negación	.848**	.498**	-	-	-	-	.790**	-	-	-	-	-
2. BIDR	.673**	.472**	.690**	-	-	-	.479**	-	.528**	-	-	-
a ² . Auto engaño	.489**	.356**	.487**	.794**	-	-	-	-	-	.828**	-	-
b ² . Manejo de la impresión	.660**	.484**	.655**	.857**	.388**	-	.608**	.345*	.623**	.708**	-	-
3. CEC	-	-	-	-	-	-	-	-	-	-	-	-

Nota. CEC = Cuestionario de Evaluación de la Capacitación. MCSDS = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación. BIDR = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. Correlaciones por instructor externo vs Evaluación favorable inducida N = 56; Correlaciones por instructor interno vs Instrucción Estándar N = 46; * p < .05, ** p < .01.

En la tabla 6 se puede observar que los nueve reactivos del CEC, muestran correlaciones significativamente más altas con la escala del BIDR, ($r = .237$, a $r = .312$, $p < .01$), mientras que las correlaciones más bajas se observan con la escala del MCSDS ($r = .139$ $p < .05$ a $r = .216$, $p < .01$).

Tabla 6
Intercorrelaciones entre los reactivos del CEC y las escalas de DS de MCSDS y BIDR

Reactivos del CEC	Correlaciones					
	1	a ¹	b ¹	2	a ²	b ²
1 ¿En qué grado demostró dominio del tema?	.211**	.231**	-	.288**	.250**	.230**
2 ¿Desarrolló el tema en forma clara y comprensible?	.139*	.158*	-	.273**	.233**	.217**
3 ¿En qué grado despertó su interés en los temas expuestos?	.210**	.175**	.121*	.303**	.241**	.270**
4 ¿Resolvió todas sus dudas en forma clara?	.204**	.157*	.126*	.312**	.249**	.260**
5 ¿En qué grado relacionó los temas con aplicaciones en el trabajo?	.177**	.162**	-	.288**	.232**	.257**
6 ¿Permitió la intervención de los participantes?	.205**	.199**	-	.287**	.229**	.261**
7 ¿Como transmitió el conocimiento en relación al Objetivo del Curso?	.216**	.202**	-	.237**	.188**	.197**
8 ¿Inició y terminó puntualmente las sesiones?	.198**	.214**	-	.268**	.245**	.206**
9 ¿Considera que se cumplió el objetivo del evento?	.191**	.173**	-	.264**	.219**	.227**

Nota. N = 263. CEC = Cuestionario de Evaluación de la Capacitación. 1 = Escala de DS de Marlowe-Crowne, a1 = subescala de atribución. b1 = subescala de negación, 2 = Balanced Inventory of Desirable Responding, a2 = Auto engaño. b2 = Manejo de la impresión. * $p < .05$, ** $p < .01$.

En la tabla 7 se muestra las intercorrelaciones de los nueve reactivos del CEC y las escalas del MCSDS y BIDR, segmentada por instrucción inducida y estándar. Se encontraron correlaciones significativamente más altas bajo instrucciones de tipo estándar con la escala del BIDR, ($r = .207$ $p < .05$, a $r = .354$, $p < .01$), mientras que las correlaciones más bajas se observan con la escala del MCSDS ($r = .195$ $p < .05$ a $r = .271$, $p < .01$).

Tabla 7
Intercorrelaciones de los reactivos del CEC, MCSDS y BIDR y sus subescalas, en función al tipo de instrucción.

Instrucción	Correlaciones											
	Evaluación favorable inducida ^a						Estándar ^b					
Reactivos del CEC/Escalas	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1	.225**	.225**	-	.233**	.196*	-	.352**	.274**	.318**	.199*	.260**	-
2	.251**	.235**	.185*	-	-	-	.295**	.232**	.252**	.195*	.253**	-
3	.297**	.218**	.284**	.194*	-	.237**	.307**	.262**	.266**	.233**	.276**	-
4	.275**	.203*	.248**	-	-	.202*	.354**	.296**	.284**	.271**	.272**	-
5	.229**	.192*	.194*	-	-	.202*	.350**	.270**	.332**	.197*	.258**	-
6	.281**	.216*	.270**	-	-	.224*	.294**	.241**	.258**	.248**	.324**	-
7	.266**	.220**	.217**	.220*	-	.232**	.207*	.158*	.185*	.213*	.278**	-
8	.243**	.182*	.222**	-	-	.205*	.292**	.305**	.194*	.217**	.306**	-
9	.257**	.210*	.222**	-	-	.189*	.278**	.230**	.251**	.239**	.298**	-

Nota. Evaluación favorable inducida^a N = 123; Estándar^b N = 140. CEC = Cuestionario de Evaluación de la Capacitación. 1 = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación, 2 = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. * $p < .05$, ** $p < .01$. La descripción de los reactivos del 1 al 9 se especifica en la tabla 6.

En lo que respecta a las intercorrelaciones de los 9 reactivos del CEC y las escalas del MCSDS y BIDR, en función del tipo de instrucción e instructor interno, la tabla 8 muestra que existe una correlación significativamente más altas con la escala del BIDR, ($r = .293$ $p < .01$, a $r = .416$, $p < .01$) con las instrucciones que inducen una evaluación favorable, mientras que la escala del MCSDS probó tener mayor relación con el CEC con las instrucciones estándar con una rango de $r = .223$ $p < .05$ a $r = .399$, $p < .01$ y el BIDR de $r = .243$ $p < .05$ a $r = .365$.

Tabla 8

Intercorrelaciones de los reactivos del CEC, MCSDS y BIDR y sus subescalas, en función al tipo de instrucción e instructor interno.

Instrucción	Correlaciones											
	Evaluación favorable inducida						Estándar					
Reactivos del CEC/Escalas	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1	.313**	.313**	-	.299*	-	-	.365**	.266**	.348**	.271**	.373**	-
2	.359**	.325**	.263*	-	-	-	.298**	-	.288**	.268**	.371**	-
3	.403**	.270*	.417**	.257*	-	.313**	.243*	-	.235*	.224*	.300**	-
4	.416**	.313**	.378**	-	-	.276*	.337**	.267**	.283**	.322**	.400**	-
5	.327**	.278*	.276*	-	-	-	.350**	.244*	.337**	.223*	.341**	-
6	.351**	.280**	.326**	-	-	.321**	.275**	.216*	.256*	.305**	.428**	-
7	.293**	.264*	.246*	-	-	.288*	-	-	-	.297**	.426**	-
8	.355**	.294**	.323**	-	-	-	.265**	.270**	-	.277**	.407**	-
9	.318**	.258*	.291**	-	-	-	.313**	.203*	.338**	.399**	.465**	-

Nota. Evaluación favorable inducida^a N = 67; Estándar^b N = 94. CEC = Cuestionario de Evaluación de la Capacitación. 1 = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación, 2 = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. * $p < .05$, ** $p < .01$. La descripción de los reactivos del 1 al 9 se especifica en la tabla 6.

Las evaluaciones hacia el instructor externo que fueron administradas con instrucciones inducidas y estándar, mostraron una nula relación con las dos escalas de DS, excepto los reactivos 3 y 4 se relacionaron con el BIDR en un rango de .304 a .317 $p < .05$ bajo instrucciones las estándar, tal como se muestra en la tabla 9.

Tabla 9

Intercorrelaciones de los reactivos del CEC, MCSDS y BIDR y sus subescalas, en función al tipo de instrucción e instructor externo.

Instrucción	Correlaciones											
	Evaluación favorable inducida						Estándar					
Reactivos del CEC/Escalas	1	a ¹	b ¹	2	a ²	b ²	1	a ¹	b ¹	2	a ²	b ²
1	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	.317*	.313*	-	-	-	-
4	-	-	-	-	-	-	.304*	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	.299*	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-	-

Nota. Evaluación favorable inducida^a N = 56; Estándar^b N = 46. CEC = Cuestionario de Evaluación de la Capacitación. 1 = Escala de DS de Marlowe-Crowne, a¹ = subescala de atribución. b¹ = subescala de negación, 2 = Balanced Inventory of Desirable Responding, a² = Auto engaño. b² = Manejo de la impresión. * $p < .05$, ** $p < .01$. La descripción de los reactivos del 1 al 9 se especifica en la tabla 6.

La tabla 10 resumen del ANOVA, contiene las fuentes de variación de los efectos presentes en el modelo constituido por el tipo de instructor, tipo de instrucción y la interacción. Los tres efectos incluidos en el modelo (tipo de instructor, tipo de instrucción, tipo de instructor*tipo de instrucción), explican menos del 1% de la varianza de las puntuaciones de la evaluación del desempeño de los instructores. Los efectos individuales de los dos factores incluidos en el modelo: *tipo de instructor* y *tipo de instrucción*, indican que no existen diferencias significativas en las puntuaciones de la evaluación del desempeño de los instructores, de manera particular. El estadístico F correspondiente a la interacción entre el *tipo de instructor* y *tipo de instrucción*, indica que no existe un efecto significativo entre dicha interacción, respecto a las puntuaciones del CEC.

Fuente	df	SS	F	Sig.
Modelo Corregido	3	38.611	1.529	.207
Intersección	1	1955902.982	77444.754	.000
Tipo de Instructor	1	65.262	2.584	.109
Tipo de Instrucción	1	.021	.001	.977
Instructor*Instrucción	1	47.382	1.876	.172
Residual	277	25.255		
Total	281			

R cuadrado = .016 (R cuadrado corregido = .006)

Mediante el procedimiento de regresión por pasos se realizaron 2 análisis a fin de conocer el mejor modelo predictor del CEC y de cada uno de los 9 reactivos que lo componen. Tanto en el primero como en el segundo análisis, las subescalas Manejo de la impresión y Autoengaño del BIDR, actuaron como variables predictoras por tener mayor correlación en los puntajes del CEC. En la tabla 11 se muestra el resultado del análisis de regresión para el modelo del CEC ($F_{9.647} = 18.351$) obteniéndose un valor de regresión múltiple igual a 0.117 el cual explica el 11% de la varianza. El modelo resultante fue Autoengaño ($t = 3.475$, $p < 0.001$) y Manejo de la impresión ($t = 3.106$, $p < 0.001$) y su ecuación $CEC = 81.93 + (.310\text{Autoengaño}) + (.252 \text{ Manejo de la impresión})$.

Variable Predictor	B	SE B	β	R^2	ΔR^2
Paso 1:				.117*	.110
Auto engaño	.310	.089	.215		
Manejo de la impresión	.252	.081	.192		

* $p < .001$.

El segundo análisis de regresión múltiple (tabla 12), fue aplicado a los nueve reactivos del CEC. El modelo principal está representado por el reactivo 1: ¿En qué grado demostró dominio del tema? (F15.53) y con una contribución ligeramente menor el reactivo 6: ¿Permitió la intervención de los participantes? (F15.46). El modelo resultante para el reactivo 1 es Manejo de la impresión ($t= 3.09$, $p < 0.05$) y Autoengaño ($t= 2.96$, $p < 0.05$). Su ecuación: $\text{Reactivo1} = 9.00 + (.035 \text{ Manejo de la impresión}) + (.036 \text{ Autoengaño})$. El coeficiente R^2 obtenido fue de 0.101, lo que explica un 10% de la varianza. El reactivo 6 queda representado de la siguiente forma: Manejo de la impresión ($t= 2.17$, $p < 0.05$) y Autoengaño ($t= 3.78$, $p < 0.001$). La ecuación resultante es: $\text{Reactivo6} = 9.15 + (.021 \text{ Manejo de la impresión}) + (.040 \text{ Autoengaño})$. El porcentaje de varianza explicado se ubicó en un 10% ($R^2 = 0.100$). Tanto el reactivo 3 ¿En qué grado despertó su interés en los temas expuestos? (F14.057), como el reactivo 4 ¿Resolvió todas sus dudas en forma clara? (F14.098), obtuvieron el mismo coeficiente $R^2 = .092$, lo que supone un 9.2% de sus varianzas. El modelo derivado del reactivo 3 es Manejo de la impresión ($t= 2.67$, $p < 0.05$) y Autoengaño ($t= 3.08$, $p < 0.05$) y la ecuación $\text{Reactivo3} = 9.02 + (.029 \text{ Manejo de la impresión}) + (.037 \text{ Autoengaño})$. El reactivo 4 presenta un modelo = Manejo de la impresión ($t= 2.47$, $p < 0.05$) y Autoengaño ($t= 3.27$, $p < 0.05$) y una ecuación del $\text{Reactivo4} = 8.97 + (.028 \text{ Manejo de la impresión}) + (.041 \text{ Autoengaño})$. El reactivo 2 ¿Desarrolló el tema en forma clara y comprensible? (F12.983), obtuvo un $R^2 = .085$, lo que explica un 8.5% de la varianza, un modelo = Manejo de la impresión ($t= 2.62$, $p < 0.05$) y Autoengaño ($t= 2.90$, $p < 0.05$) y finalmente su ecuación resultante fue $\text{Reactivo2} = 9.09 + (.028 \text{ Manejo de la impresión}) + (.034 \text{ Autoengaño})$. El reactivo que aborda el grado en que el instructor relacionó los temas con aplicaciones en el trabajo (R5), con una aportación ligeramente menor a la del reactivo 2, obtuvo un coeficiente $R^2 = .082$, es decir un 8.2% de la varianza. El modelo y la ecuación de éste reactivo quedo representado de la siguiente forma: Manejo de la impresión ($t= 2.55$, $p < 0.05$) y Autoengaño ($t= 2.86$, $p < 0.05$) y $\text{Reactivo5} = 9.09 + (.028 \text{ Manejo de la impresión}) + (.034 \text{ Autoengaño})$, respectivamente (F12.435).

Variable Predictor	B	SE B	β	R^2	ΔR^2
Reactivo 1				.101*	.094
Manejo de la impresión	.035	.011	.193		
Auto engaño	.036	.012	.185		
Reactivo 2				.085*	.079
Manejo de la impresión	.028	.011	.165		
Auto engaño	.034	.012	.183		
Reactivo 3				.092*	.085
Manejo de la impresión	.029	.011	.167		
Auto engaño	.037	.012	.193		
Reactivo 4				.092*	.086
Manejo de la impresión	.028	.011	.155		
Auto engaño	.041	.013	.205		
Reactivo 5				.082*	.076
Manejo de la impresión	.028	.011	.161		
Auto engaño	.034	.012	.181		
Reactivo 6				.100*	.094
Manejo de la impresión	.021	.010	.135		
Auto engaño	.040	.011	.236		

* $p < .001$. La descripción de los reactivos del 1 al 9 se especifica en la tabla 6.

CAPITULO V DISCUSION Y CONCLUSIONES

Todas las organizaciones necesitan conocer su situación y las condiciones en que se encuentran, así como las alternativas que ofrecen para la solución de problemas de la sociedad. Un diagnóstico organizacional contribuye a determinar cuál es la situación de la organización, cómo se relacionan los individuos y los grupos, y qué potenciales pueden ser explotadas, así como cuales son las dificultades que enfrenta el sistema organizacional (D. Rodríguez, 1999).

Se reconoce ampliamente que no existe una evaluación objetiva de los resultados de la capacitación en las organizaciones, debido a que no se cuenta con instrumentos válidos y confiables para tal fin. Se destinan grandes presupuestos en la capacitación, en especial en sector de la industria petrolera en México, sin embargo, no existe en la mayoría de los casos una evaluación de los resultados de capacitación.

La distorsión y el sesgo en las respuestas a los cuestionarios de evaluación de la capacitación, no ha sido debidamente estudiada. En este sentido los estudios sobre la DS se han enfocado únicamente a la evaluación de la personalidad y aunque algunos autores han considerado que pudiera haber cierta influencia de la DS en las distorsión de las respuestas sobre la evaluación de la capacitación, aun no hay estudios que lo comprueben (D. L. Kirkpatrick, 2000).

El propósito del presente estudio fue el obtener evidencia del efecto de las instrucciones socialmente deseables en las respuestas de 281 participantes a cursos de capacitación, quienes evaluaron el desempeño de 28 instructores (externos e internos), en

eventos de tipo administrativo, desarrollo humano, informático, seguridad y técnicos en una empresa paraestatal en México.

A pesar del hecho de que las escalas de MCSDS y BIDR fueron desarrolladas para medir la tendencia de las personas a distorsionar autoreportes que miden actitudes y comportamientos socialmente deseables, se comprobó que existe una relación significativamente positiva, aunque moderada, en el CEC y las escalas de DS. Lo que indica que una proporción de DS se presenta en las estimaciones de las evaluaciones de la capacitación. Por lo tanto puede pensarse que cuando una personas evalúa el desempeño de un instructor, de tal manera que las imágenes de estos se vean implicadas por sus respuestas, los reactivos del CEC serán contestados por presiones normativas más que por querer proporcionar una evaluación verídica (Li & Bagger, 2007).

En un primer análisis los 9 ítems del que consta el CEC, mostraron correlaciones significativamente más altas, con las escala de DS del BIDR. En una segunda fase en la que se segmentaron las instrucciones inducidas y estándar, se encontraron correlaciones significativamente más altas bajo instrucciones de tipo estándar con la escala del BIDR que con la escala del MCSDS y el reactivo 4 que mide la capacidad que tiene el instructor en resolver todas las dudas de los participantes en forma clara fue una alternativa más deseable cuando se siguieron instrucciones estándar. Sin embargo al analizar los resultados ahora por tipo de instrucción y tipo de instructor se observó que únicamente los instructores internos cuando son evaluados con instrucciones inducidas se relacionan más significativamente con la escala del BIDR, y nuevamente el reactivo 4 alcanzó un relación significativamente más alta (.400). Los reactivos que miden la capacidad que tiene el instructor en resolver todas

las dudas de los participantes en forma clara, así como los reactivos que miden la habilidad de transmitir los conocimientos del instructor en relación al objetivo del curso, mostraron ser una alternativa de respuesta en los reactivos del CEC, socialmente más convenientes o deseables que otras. Lo que puede pensarse entonces, que las respuestas que expresan las capacidades pedagógicas de los instructores, están determinadas por la conveniencia social de éstas (Braun et al., 2002; Kerlinger & Howard, 2002; Nunnally & Bernstein, 1995).

En lo que concierne a las instrucciones administradas en este estudio, se esperaba que la instrucción que inducía evaluar favorablemente al instructor, mostrara una relación significativamente mayor con las escalas de MCSDS y BIDR, sin embargo la instrucción estándar mostró una correlación ligeramente más alta ($r = .389$, $p < .01$). Esto puede explicarse en parte, debido a que en dichas instrucciones se especifica que la información proporcionada será tratada de manera confidencial, pero no anónima ya que el trabajador debe poner sus datos personales. Esta situación para los respondientes, en ocasiones, es interpretada como una situación amenazante ya que ven “comprometida” su situación laboral a partir de lo que puedan o no responder, por lo que pueden distorsionar intencionalmente lo que se les pregunta (Cabañeros et al., 2003).

Sin embargo al contrastar las variables CEC y las escalas del BIDR y MCSDS por tipo de instructor interno y el tipo de instrucción, los puntajes mostraron una correlación significativamente más alta con la escala del BIDR cuando se aplica en forma conjunta con la instrucción inducida ($r = .456$, $p < .01$), que cuando se administra con la instrucción estándar ($r = .372$, $p < .01$). Esto se puede explicar debido a que entre los participantes y el instructor interno se guarda una estrecha camaradería o “contrato psicológico”, vínculo que

no se establece con los instructores externos, adoptando así respuestas socialmente deseables con sus pares (Davis & Newstrom, 1991; Muchinsky, 2007; Palací, 2005; Rüssel, 1963; Schein, 1982). Estos resultados demuestran que las instrucciones para evaluar cursos de capacitación pueden ser deliberadamente manipulados con el objetivo de influir en las opiniones hacía un tercero (Kuncel & Tellegen, 2009) y que no solo éste fenómeno se observa en estudios en donde se aplican instrucciones inducidas en autoreportes (R. A. Gordon, 1987; Johnsen, Tracy, & Hohn, 1983; Rogers et al., 2002).

Se comprobó que las instrucciones inducidas no difieren en los promedios del CEC, respecto a las instrucciones estándar. De igual forma no existen diferencias en los puntajes del desempeño de un instructor interno en comparación a un externo, en los cursos de capacitación. Sin embargo a pesar de no encontrarse diferencias en los promedios dichas condiciones, no debe perderse de vista que en esta muestra se obtuvieron calificaciones positivas hacia los instructores internos, resultados que deberían ser cuestionados en cuanto a su validez, más allá de los errores de medición, tales como el efecto halo, la aquiescencia y el error de indulgencia (Chang & Ferre, 1988; Phelps et al., 1986; Ray, 2001; H. C. Smith & Wakeley, 1988; Thomas et al., 1979), por lo que es recomendable considerar a la DS como un estilo de respuesta en los cuestionarios que evalúan el desempeño de los instructores de capacitación.

El análisis de regresión múltiple reveló que el BIDR tiene un 11% de poder predictivo en la varianza de los puntajes del CEC. Por otra parte el análisis de regresión no mostró ninguna relación entre el MCSDS y el CEC. Los puntajes que reflejan el nivel mostrado por el instructor en cuanto al dominio del tema (reactivo 1), así como los puntajes

que obtienen los instructores del grado en que estos permiten la intervención de los participantes (reactivo 6), puede predecirse hasta en 10% por la DS. Parece ser, que de la misma forma (consciente o inconscientemente) en la que manipulamos nuestras respuestas al dar una buena impresión de nuestras capacidades personales, cuando se trata de obtener algún beneficio en particular, también se manipula las respuestas cuando se reportar las capacidades de un tercero (en este caso conocimientos temáticos y habilidades pedagógicas), es decir se considera que mi evaluación hacia otra persona en sí será una respuestas socialmente deseable, lo que implica entonces sobre dimensionar sus conductas y otorgarle indirectamente un beneficio. Por otra parte reportar el grado en que el instructor despierta el interés del participante en los temas expuestos, así como el responder si resolvió todas sus dudas en forma clara (nuevamente), pueden ser preguntas comprometidas para el respondiente. Este efecto es predicho en un 9% por la DS.

A pesar del hecho de que la escala de BIDR fue desarrollada para medir la tendencia de las personas a distorsionar autoreportes que miden actitudes y comportamientos socialmente deseables, los resultados de esta investigación demuestra ser una herramienta eficaz para la validez del CEC, sin olvidar que solo aplica para la población participante para este estudio.

Un importante punto a considerar, es el contemplar que los cuestionarios no fueron anónimos, dado los requerimientos internos de la organización en la que se llevó a cabo el estudio. Esto hace suponer que el reconocer el nombre de quien reporta, puede generar cierta disposición a distorsionar la respuesta, por lo que es recomendable realizar en subsiguientes estudios la aplicación de los instrumentos cuidando el anonimato de los

encuestados, a fin de comparar los resultados con el presente estudio. De igual forma a fin de garantizar que las instrucciones se lleven a cabo con estricto apego, es recomendable que se lean en voz alta a todos los participantes, lo que implica trabajar con grupos independientes por tipo de instrucción.

Se recomienda correr este mismo procedimiento en un trabajo posterior con un grupo experimental y un grupo control. En el grupo experimental se le presentaran dos tipos de instructores en diferentes tiempos; uno con un nivel de competencia alto y el otro bajo, impartiendo el mismo tema. En el grupo control 1 dos instructores con un rendimiento alto y en el grupo control 2, dos instructores con un rendimiento bajo, ambos instructores presentados en diferentes tiempos durante el curso. El propósito radica en medir el grado DS, en diferentes niveles de desempeño, situación que puede considerarse como una limitante en este estudio. Asimismo se sugiere realizar cambios sustanciales al cuestionario, como por ejemplo: proporcionar instrucciones claras eliminar el nombre del participante, cambiar el tipo de respuesta de tipo Likert a un formato dicotómico de “Si” y “No”, mantener la redacción de los reactivos tan simple como sea posible, evitar realizar la aplicación en el aula y en vez de ello enviarlo por correo o asistir al lugar del trabajador, a fin de evitar el consenso con sus demás compañeros o bien puede diseñarse un cuestionario programado y asistido por computadora. Queda claro que hay mucho que trabajar en el diseño, elaboración y análisis de los reactivos que conforman un cuestionario que evalúa el desempeño de los instructores y que habría que dedicarle un apartado más extenso a este respecto. Sin pretender resolver este conflicto y considerando que no es el objetivo de este

estudio, se propone un cuestionario alternativo que sirva de punto de partida para la mejora en las medidas del CEC (ver anexo E).

No hay que olvidar que la DS es la fuente que más afecta la validez de los experimentos en la psicología y las ciencias sociales (King & Bruner, 2000), así que los resultados encontrados en las correlaciones en el CEC y las escalas del MCSDS y BIDR, sugieren seriamente que los reactivos que evalúan el desempeño de instructores de capacitación, particularmente en este estudio, pueden estar contaminados por éste estilo de respuesta, lo que hace suponer que las evaluaciones de cursos de capacitación aquí evaluados, son susceptibles de distorsión y pueden dar como resultado serios problemas en la validez de los datos, por lo que no se descartar la presencia de la DS en los resultados de la evaluación de la capacitación, aún y cuando la DS se manifieste exclusivamente en los autoreportes que miden actitudes y comportamientos socialmente deseables.

Cada evaluación que se realiza durante el proceso de capacitación, está encaminada a obtener información de la percepción individual de los participantes acerca del programa de capacitación y así mismo detectar el nivel de conocimiento que están alcanzando los capacitandos; sin embargo, el no tener evidencia de que los participantes muestran a sus instructores de una manera demasiado positiva, impide determinar si se logró el objetivo de la capacitación y por ende, realizar los ajustes necesarios a los contenidos temáticos de los cursos, así como determinar cuál es el ritmo más adecuado con el que se deben desarrollar los temas. Al no contar con evaluaciones confiables, el encargado de planificar la capacitación no se encuentra en posibilidad de demostrar que los cursos han alcanzado los

objetivos esperados para el desarrollo de los conocimientos, habilidades y actitudes del personal (Grados, 2001).

Al contabilizar cuestionarios de evaluación de la capacitación distorsionados de manera positiva, se sobredimensiona un programa de capacitación. Esto conlleva a seguir programando la intervención de los mismos instructores o de la continuidad de determinado evento, en vez de realizar los ajustes necesarios o cambios sustanciales más pertinentes. Las consecuencias de una deficiente detección de la eficacia de un curso de capacitación pueden reflejarse drásticamente en la incorrecta formación del personal, el óptimo desarrollo de las operaciones y, por ende, en un impacto negativo para el logro de los objetivos organizacionales.

Sin duda uno de los elementos trascendentales que se contempla en la capacitación, es evaluar en forma integral el sistema de capacitación implantado. Diversos autores han ponderado a la evaluación de los resultados de la capacitación, cómo una de las fases más importantes de los programas de capacitación en las organizaciones y de igual forma han realizado diversos esfuerzos por desarrollar y aplicar nuevas metodologías que proporcionen validez en sus resultados (Blignaut, 1993; Ford, 1984; D. L. Kirkpatrick & Kirkpatrick, 2007; Kumpikaitė, 2007; Toplis, 1992). La importancia de la evaluación radica en medir el impacto de la capacitación en el escenario laboral, lo que da como resultado el costo beneficio de un curso, en función de la manera que los capacitandos aplican los conocimientos obtenidos. Sin embargo y a pesar de que la evaluación de los resultados de la capacitación debería ser metódica, sistematizada y jerarquizada con un enfoque actualizado y vanguardista, confiable, válida, consistente y objetiva dentro de los proceso

de la organización (Llanos, 2005), hoy en día sigue siendo una de las áreas menos investigadas y descuidadas en la psicología organizacional .

Las organizaciones en general no han recibido los beneficios a corto y mediano plazo que tanto promete la psicología en sus propuestas teóricas, porque sus directivos no ven certidumbre en la aplicación de la metodología psicológica con la que se pueda describir, explicar y predecir los resultados esperados de los programas de capacitación. En un ambiente de censura, incredulidad y miopía de los altos ejecutivos, en donde no se permite invertir en el tiempo necesario para diseñar e implantar instrumentos para medir los resultados de la capacitación, se seguirán aplicando los tradicionales procedimientos desordenados, así como los mismos recursos inadecuados, carentes de validez y confiabilidad limitando así su desarrollo y consolidación (Pinto, 2000). Actualmente las organizaciones no conocen los rendimientos que se pueden obtener al aplicar las técnicas de investigación-acción, con las cuales la evaluación del adiestramiento podría tomar forma de un experimento, tomando medidas de eficiencia de “antes” y “después” en grupos control (no adiestrados) y experimental (adiestrados) (Bunker & Cohen, 1977; Dunnette & Kirchner, 1989; Guzmán & Sánchez Escobedo, 2006), con las que se pueden identificar áreas de oportunidad para el desarrollo del capital humano y determinar acciones estratégicas de capacitación para su solución. Aun estamos en espera de que se le acredite a la psicología organizacional como una disciplina capaz de generar datos esenciales que contribuyen en la solución de los problemas de capacitación.

La aplicación de la escala de DS del BIDR, administrada a una muestra de trabajadores de una empresa paraestatal, ofrece una nueva alternativa para medir la validez

de los cuestionarios que evalúan el desempeño de los instructores de capacitación en el entorno de las organizaciones. Este instrumento puede utilizarse como un soporte a los formatos de respuesta tradicionalmente utilizados para evaluar el desempeño de los instructores. Con su aplicación se está en la posibilidad de controlar las respuestas que sobre dimensionan los resultados del ejercicio de los instructores dentro de las aulas. Es un instrumento de fácil aplicación, que puede ser aplicado en un tiempo reducido. Los resultados de esta investigación respaldan el uso de las escalas de DS para estimar sesgos de respuestas en un sentido socialmente deseable, al utilizar reportes tipos CEC. Con estos hallazgos se confirma la presencia de la DS como una fuente de error en las evaluaciones a instructores y que solo era atribuible a los autoreportes de personalidad. En definitiva el presente trabajo asume la importancia de que las medidas de personalidad son una herramienta efectiva en la toma de decisiones para los procesos de formación en las organizaciones. De este estudio se desprende la necesidad de modificar los sistemas de evaluación de la capacitación, desde adoptar las estrategias más populares como el informar a los trabajadores de que sus respuestas distorsionadas podrá tener consecuencias negativas para los instructores, hasta la administración sistemática de las escalas de DS en forma alterna con los instrumentos que se aplican para la recogida de datos al término de los cursos, correlacionarlos entre si y descartar aquellos reportes que puntúen más alto en DS. El propósito es ofrecer a las organizaciones sistemas más efectivos de medición de los resultados de los cursos de capacitación, con los cuales se reduzcan y neutralice la DS.

REFERENCIAS

- Aiken, L. R. (2003). *Test Psicológicos y Evaluación* (Undécima ed.). México: Pearson Prentice Hall.
- Anastasi, A., & Urbina, S. (1998). *Test Psicológicos* (7a. ed.). México: Prentice Hall.
- Armitage, C., & Conner, M. (1999). Predictive validity of the theory of planned behaviour: the role of questionnaire format and social desirability. *Journal of Community & Applied Social Psychology*, 9(4), 261-272.
- Avery, Z., & Gary, J. S. (1985). Cross-Validation of a short form of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 41(2), 236-238.
- Barger, S. D. (2002). The Marlowe–Crowne Affair: Short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, 79(2), 286–305.
- Bartlett, C. J., & Richard, D. (1967). Social desirability response differences under research, simulated selection, and faking instructional sets. *Personnel Psychology*, , 20(3), 281-288.
- Bentley, T. (1993). *Capacitación Empresarial*. Colombia: McGraw-Hill.
- Benzein, E., & Berg, A. (2003). The Swedish version of Herth Hope Index – an instrument for palliative care. *Scandinavian Journal of Caring Sciences*, 17 (4), 409-415.

- Beretvas, S. N., Meyer, J. L., & Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement, 62*(4), 570-589.
- Bernreuter, R.G. (1933). Validity of the personality inventory. *Personality Journal, 11*, 383-386.
- Birenbaum, M., & Montag, I. (1989). Style and substance in social desirability scales. *European Journal of Personality, 3* (1), 47-59, 13p.
- Blake, O. (2000). *Origen, Detección y Análisis de las Necesidades de Capacitación*. Argentina: Macchi.
- Blignaut, C. J. H. (1993). Towards the estandardization of an instrument for the evaluation of training. *Journal of Industrial Psychology, 19*(1), 11-14.
- Braun, H. I., Jackson, D. N., & Wiley, D. E. (2002). Socially Desirable Responding: The evolution of a construct. *En: Psychological and Educational Measurement, 49-69*.
- Bryant., S. E., & Terborg., J. R. (2008). Impact of peer mentor training on creating and sharing organizational knowledge. *Journal of Managerial Issues, XX*(1), 11-29.
- Bunker, K. A., & Cohen, S. L. (1977). The rigors of training evaluation: A discussion and field demonstration. *Personnel Psychology, 30*, 525-541.
- Cabañeros, J., García Cueto, E., & Lozano, L. M. (2003). *Efecto de la deseabilidad social en preguntas sobre temas comprometidos* Paper presented at the IX Conferencia Española de Biometría.

- Collazo, A. A. (2005). Translation of the Marlowe-Crowne Social Desirability Scale into an equivalent spanish version. *Educational and Pshychological Measurement*, 65(5), 780-806.
- Coq, D., & Asián, R. (2002). Estudio de la “deseabilidad social” en una investigación mediante encuestas a empresarios andaluces. *Metodología de encuestas*, 4(2), 211-225.
- Chang, M. K., & Ferre, V. (1988). An analysis of student teaching evaluations by college supervisors and cooperating teachers. *Education*, 108(4), 493-496.
- Crowne, D. P. y Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Darby, J. A. (2007). Are course evaluations subject to a halo effect? *Research in Education*, 77, 46-55.
- Davis, K., & Newstrom, J. W. (1991). *Comportamiento humano en el trabajo. Comportamiento Organizacional*. México: McGraw-Hill.
- DeWine, S. (1987). Evaluation of organizational communication competency. *Journal of Applied Communication Reserch*, 15(1-2), 113-127.
- Domínguez, E. A., Salas, M. I., & Reyes-Lagunes, I. (2008). Validez concurrente de la Escala de Deseabilidad Social de Domínguez utilizando la Escala de Deseabilidad Social de Marlowe-Crowne. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 25(1), 125-139.

- Dunnette, M. D., & Kirchner, W. K. (1989). *Psicología Industrial*. México: Trilas.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Ellis, L., Burke, D. M., Lomire, P., & McCormack, D. R. (2003). Student grades and average ratings of instructional quality: The need for adjustment. *Journal of Educational Research*, 97(1), 35-40.
- Ferrando, P. J., & Chico, E. (2000). Adaptación y análisis psicométrico de la escala de deseabilidad social de Marlowe y Crowne. *Psicothema* 12(3), 383-389.
- Ferrando, P. J., Lorenzo-Seva, U., & Chico, E. (2009). A General Factor-Analytic Procedure for Assessing Response Bias in Questionnaire Measures. *Structural Equation Modeling*, 16(2), 364-381.
- Fineman, S. (1975). The Work Preference Questionnaire: a measure of managerial need for achievement. *Journal of Occupational Psychology*, 48(1), 11-32.
- Ford, J. K. (1984). Introducing new methods for conducting training evaluation and for linking training evaluation to program redesign. *Personnel Psychology*, 37, 651-665.
- Fraser, J. (2007). Developing a framework to evaluate training programs provided by WHO: The feasibility of incorporating social justice, cultural competency and return on investment. *The international journal of learning*, 14(9), 103-109.

- González, H., & Iruarrizaga, I. (2005). Evaluación de las distorsiones de respuesta mediante el MMPI-2. *Papeles del Psicólogos, Consejo General de Colegios Oficiales de Psicólogos, España, 26(092)*, 129-137.
- Gordon, M. E., & Petty, M. M. (1971). A note on the effectiveness of research conditions in reducing the magnitude of dissimulation on a self-report criterion. *Personnel Psychology, 24(1)*, 53-61.
- Gordon, R. A. (1987). Social Desirability Bias: A demonstration and technique for Its reduction. *Teaching of Psychology, 14(1)*, 40-42.
- Grados, J. A. (1998). *La Capacitación*. México: Trillas.
- Grados, J. A. (2001). *Capacitación y Desarrollo de Personal*. México: Trillas.
- Guzmán, S., & Sánchez Escobedo, P. (2006). Efectos de un programa de capacitación de profesores en el desarrollo de habilidades de pensamiento crítico en estudiantes universitarios en el Sureste de México. *Revista Electrónica de Investigación Educativa, 8(2)*, 1-17.
- Harry, J. M. (1984). A revised Measure of Approval Motivation and Its relationship to social desirability. *Journal of personality assessment, 48(5)*, 508-519.
- Hartley, S. L., & MacLean Jr., W. E. (2006). A review of the reliability and validity of Likert-type scales for people with intellectual disability. *Journal of Intellectual Disability Research, 50(11)*, 813-827.

- Johnsen, E. P., Tracy, D. B., & Hohn, R. L. (1983). Induced response bias on the state-trait anxiety inventory. *Social Behavior & Personality: An International Journal*, *11* (1), 113.
- Johnson, T. P., & Fendrich, M. (2002). *A Validation Of The Crowne-Marlowe Social Desirability Scale*. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg FL.
- Kashy, D., & DePaulo, B. (1996). Who lies? *Journal of Personality and Social Psychology*, *70*, 1037-1051.
- Kerlinger, F. N., & Howard, B. L. (2002). *Investigación del Comportamiento. Métodos de Investigación en Ciencias Sociales* (Cuarta ed.). México: McGraw-Hill.
- King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, *17*(2), 79-103.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2007). *Implementing the Four Levels*. San Francisco: Berrett-Koehler.
- Kirkpatrick, D. L. (2000). *Evaluación de Acciones Formativas. Los Cuatro Niveles*. España: Epice.
- Konstabel, K., Aavik, T., & Allik, J. (2006). Social desirability and consensual validity of personality traits. *European Journal of Personality*, *20*, 549–566.
- Kumpikaitė, V. (2007). Human Resource Training Evaluation. *Engineering Economics*, *5*(55), 29-36.

- Kuncel, N. R., & Tellegen, A. (2009). A conceptual and empirical reexamination of the measurement of the social desirability of items: implications for detecting desirable response style and scale development. *Personnel Psychology*, *62*, 201–228.
- Landy, F. J., & Conte, J. M. (2005). *Psicología Industrial* (1a ed.). México: McGraw-Hill.
- Larsen, K. S. (1994). Attitudes toward the transportation of nuclear waste: The development of a Likert-Type Scale. *Journal of Social Psychology*, *134*(1), 27-34.
- Larsen, K. S., Reed, M., & Hoffman, S. (1980). Attitudes of heterosexuals toward homosexuality: A Likert-Type Scale and construct validity. *Journal of Sex Research*, *16*(3), 245-257.
- Leite, W. L., & Beretvas, S. N. (2005). Validation of Scores on the Marlowe-Crowne Social Desirability Scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, *65*(140-154).
- Lewis, C. (2000). The religiosity-psychoticism relationship and the two factors of social desirability: a response to Michael W. Eysenck (1999). *Mental Health, Religion & Culture*, *3*(1), 39-45.
- Li, A., & Bagger, J. (2007). The balance inventory of desirable responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement*, *67*(525), 525-538.
- Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *The Journal of Social Psychology*, *140*(5), 628-635.

- Llanos, J. (2005). *Integración de Recursos Humanos*. México: Trillas.
- Martínez Selva, J. (2005). *La psicología de la Mentira*. Barcelona: Paidós.
- McGehee, W., & Thayer, P. W. (1986). *Capacitación: Adiestramiento y Formación Profesional*. México: Limusa.
- Meerkerk, G.-J., Van Den Eijnden, R. J. J. M., Vermulst, A. A., & Garretsen, H. F. L. (2009). The Compulsive Internet Use Scale (CIUS): Some psychometric properties. *CyberPsychology & Behavior*, 12 (1), 1-6.
- Mills, J., Loza, W., & Kroner, D. (2003). Predictive validity despite social desirability: evidence for the robustness of self-report among offenders. *Criminal Behaviour & Mental Health* 13(2), 140.
- Muchinsky, P. M. (2007). *Psicología aplicada al trabajo* (8va. ed.). México: CENGAGE Learning.
- Nunnally, J. C., & Bernstein, I. J. (1995). *Teoría Psicométrica* (Tercera ed.). México: McGraw-Hill.
- O'Grady, K. E. (1988). The Marlowe-Crowne and Edward Social Desirability Scale: A Psychometric Perspective. *Multivariate Behavioral Research*, 23, 87-101.
- Ommundsen, R., & Larsen, K. S. (1997). Attitudes toward illegal aliens: The reliability and validity of a Likert-Type Scale. *Journal of Social Psychology*, 137 (5), 665-667.
- Palací, D. F. J. (2005). *Psicología de la Organización*. Madrid: Prentice-Hall.

- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measure of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR Version 6*. Unpublished manuscript, Vancouver, Canada.
- Paulhus, D. L., Fridhandler, B., & Hayes, S. (1997). Psychological Defense, Contemporary Theory And Research. In A. Press (Ed.), *Handbook of Personality Psychology* (pp. 543-561). San Diego
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and Denial in Socially Desirable Responding. *Journal of Personality and Social Psychology*, *60*(2), 307-317.
- Phelps, L., Schmitz, C. D., & Boatright, B. (1986). The Effects of Halo and leniency on cooperating teacher reports using Likert-Type Rating Scales. *Journal of Educational Research*, *79*(3), 151-154.
- Pinto, V. (2000). *Planeación Estratégica de la Capacitación*. México: McGraw-Hill.
- Puccio, G. J., Firestien, R. L., Coyle, C., & Masucci, C. (2006). A review of the effectiveness of CPS training: A focus on workplace issues. *Journal compilation*, *15*(1), 19-33.

- Ramanaiah, N. V., & Martin, H. J. (1980). On the two-dimensional nature of the Marlowe-Crowne Social Desirability Scale. *Journal of Personality Assessment, 44*(5), 505-514.
- Ray, J. J. (1984). The reliability of short social desirability scale. *The Journal of Social Psychology, 123*, 133-134.
- Ray, J. J. (2001). Acquiescence and problems with forced choice scales. *The Journal of Social Psychology, 130*(3), 397-399.
- Reavis, C. A. (1979). A study of the effects of prefatory remarks on teacher evaluation. *Journal of Educational Research, 72* (3), 173-179.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38* (1), 119-125.
- Rodríguez, D. (1999). *Diagnóstico Organizacional*. México: Alfaomega.
- Rodríguez, E. (1993). *Administración de la Capacitación*. México: McGraw-Hill.
- Rogers, R., Vitacco, M. J., Jackson, R. L., Martin, M., Collins, M., & Sewell, K. W. (2002). Faking psychopathy? an examination of response styles with antisocial youth. *Journal of Personality Assessment, 78* (1), 31-46.
- Rüssel, A. (1963). *Psicología del trabajo*. Madrid: Ediciones Morata.
- Russell, P. A., Lankford, M. W., & Grinnell Jr., R. M. (1981). A Management Evaluation Scale for social workers. *Journal of Psychology, 107* (1), 127-130.

- Salgado, J. F. (2005). Personalidad y deseabilidad social en contextos organizacionales: implicaciones para la práctica de la psicología del trabajo y las organizaciones. *Papeles del Psicólogos*, 26, 115-128.
- Schein, E. H. (1982). *Psicología de la Organización* (3a ed.). México: Prentice Hall.
- Schweitzer, R., Perkoulidis, S., Krome, S., Ludlow, C., & Ryan, M. (2005). Attitudes towards refugees: The dark side of prejudice in Australia. *Australian Journal of Psychology*, 57(3), 170-179.
- Seol, H. (2007). A Psychometric Investigation of the Mariowe-Crowne Social Desirability Scale Using Rasch Measurement. *En: Measurement and Evaluation in Counseling and Development*, 40, 155-168.
- Siliceo, A. (2001). *Capacitación y desarrollo de personal*. México: Limusa.
- Silverthorn, N. A., & Gekoski, W. L. (1995). Social desirability effects of measures of adjustment to university, independence from parents. *Journal of Clinical Psychology*, 51(2), 244-251.
- Smith, H. C., & Wakeley, J. H. (1988). *Psicología de la Conducta Industrial*. México: McGraw-Hill.
- Smith, S., Smith, K., & Seymour, K. (1993). Social Desirability of Personality Items as a Predictor of Endorsement: A Cross-Cultural Analysis. *Journal of Social Psychology*, 133(1), 43-52.

- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17). *European Journal of Psychological Assessment, 17*(3), 222-232.
- Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment, 78* (2), 370-389.
- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 28*(2), 191-193.
- Templer, D. I., King, F. L., Brooner, R. K., & Corgiat, M. (1984). Assessment of body elimination attitude. *Journal of Clinical Psychology, 40*(3), 754-759.
- Thomas, C. B., Hall, J. A., Miller, F. D., Dewhirst, J. R., Fine, G. A., Taylor, M., et al. (1979). Evaluation, apprehension, social desirability, and the interpretation correlations. *Social Behavior and Personality, 7*(2), 193-197.
- Toplis, J. (1992). Training evaluation-reflections on the first steps. *European Work & Organizational Psychologist, 2*(4), 354-367.
- Van de Vijver, F. J. R. (1997). Bias research: Introduction. *European Review of Applied Psychology, 47*(4), 261-262.
- Van de Vijver, F. J. R. (1998). Towards a theory of bias and equivalence. *ZUMA-Nachrichten Spezial, 3*, 41-65.
- Zepeda, F. (1999). *Psicología Organizacional*. México: Addison Wesley.

Zook II, A., & Sipps, G. J. (1985). Cross-validation of a short form of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 41*(2), 236-238.

ANEXO A

INSTRUCCIONES PARA EL LLENADO DEL CEC

1. Evaluación favorable inducida (Instructor Interno).

Usted conoce los beneficios a la empresa al contar con instructores internos, por lo que queremos dar una buena impresión a nuestra máxima representación sindical y a los directivos de la actual administración, sobre el desarrollo del curso que acaba usted de concluir.

Queremos convencerlos del excelente desempeño del instructor, apoyados de su valiosa aportación, sumándola a la de todos los participantes que han asistido antes que usted.

Le agradecemos, de antemano su apoyo para el llenado de esta evaluación.

2. Evaluación favorable inducida (Instructor Externo).

Usted conoce los beneficios a la empresa al contratar instructores externos, por lo que queremos dar una buena impresión a nuestra máxima representación sindical y a los directivos de la actual administración, sobre el desarrollo del curso que acaba usted de concluir.

Queremos convencerlos del excelente desempeño del instructor, apoyados de su valiosa aportación, sumándola a la de todos los participantes que han asistido antes que usted.

Le agradecemos, de antemano su apoyo para el llenado de esta evaluación.

3. Instrucciones Estándar para el Instructor Interno y Externo.

Para mejorar la calidad de la capacitación es necesaria su valiosa colaboración. Por favor responda cada pregunta cruzando el cuadro que mejor exprese su opinión.

La información es confidencial y es de uso exclusivo para el área de capacitación.

ANEXO B

CUESTIONARIO DE EVALUACIÓN DE LA CAPACITACIÓN (CEC)

Evaluación Final del Evento por los Participantes

		Para uso exclusivo de Recursos Humanos					
Nombre del Evento		Clave del evento					
Ficha del participante		Área de Adscripción					
Nombre del Instructor		Fechas:		Inicio		Término	
INSTRUCTOR		No aplica	Malo	Regular	Bueno	Muy Bueno	Excelente
1.- ¿En que grado demostró dominio del tema ?	NA	6	7	8	9	10	
2.- ¿Desarrolló el tema en forma clara y comprensible ?	NA	6	7	8	9	10	
3.- ¿En que grado despertó su interés en los temas expuestos ?	NA	6	7	8	9	10	
4.- ¿Resolvió todas sus dudas en forma clara ?	NA	6	7	8	9	10	
5.- ¿En que grado relacionó los temas con aplicaciones en el trabajo ?	NA	6	7	8	9	10	
6.- ¿Permitió la intervención de los participantes ?	NA	6	7	8	9	10	
7.- ¿Como transmitió el conocimiento en relación al Objetivo del Curso ?	NA	6	7	8	9	10	
9.- ¿Inició y terminó puntualmente las sesiones?	NA	6	7	8	9	10	
10.- ¿Considera que cumplió el objetivo del evento?	NA	6	7	8	9	10	

ANEXO C

INVENTARIO EQUILIBRADO DE RESPUESTA DESEABLE (BIDR)

PDS (BIDR Versión 7)

Nombre: _____ Edad: _____ Género: Masculino Femenino

Fecha: ____/____/____
Mes Día Año

Instrucciones: Lea cada enunciado, y circule el número de la opción que mejor lo describa, desde *No es Cierto* hasta *Muy Cierto* sobre usted.

	No es Cierto				Muy Cierto
1. Mis primeras impresiones sobre la gente usualmente...	1	2	3	4	5
2. Me resultaría difícil romper alguno de mis ...	1	2	3	4	5
3. No me importa saber lo que las otras personas realmente...	1	2	3	4	5
4. No siempre he sido honesto...	1	2	3	4	5
5. Siempre se porqué me agradan ...	1	2	3	4	5
6. Cuando mis emociones son intensas...	1	2	3	4	5
7. Una vez que me he decidido, otras personas no pueden ...	1	2	3	4	5
8. No soy un conductor seguro cuando excedo ...	1	2	3	4	5
9. Estoy enteramente en control de mi ...	1	2	3	4	5
10. Es difícil para mi callar un ...	1	2	3	4	5
11. Nunca me arrepiento de ...	1	2	3	4	5
12. En ocasiones me pierdo de algunas cosas porque no ...	1	2	3	4	5
13. La razón por la que voto es porque mi voto puede ...	1	2	3	4	5
14. La gente no me nota a mi...	1	2	3	4	5
15. Soy una persona completamente ...	1	2	3	4	5
16. Rara vez acepto ...	1	2	3	4	5
17. Soy muy confiado en ...	1	2	3	4	5
18. En ocasiones he dudado de mi habilidad ...	1	2	3	4	5
19. Está bien por mi si a algunas personas ...	1	2	3	4	5
20. Soy solamente una persona ...	1	2	3	4	5
21. En ocasiones digo mentiras si tengo ...	1	2	3	4	5
22. Nunca cubro mis ...	1	2	3	4	5
23. Ha habido ocasiones en las que he tomado ...	1	2	3	4	5
24. Nunca digo ...	1	2	3	4	5
25. En ocasiones trato de quedar a mano más que ...	1	2	3	4	5
26. Siempre obedezco las leyes, aunque sea improbable...	1	2	3	4	5
27. He dicho algo malo sobre algún amigo o amiga a ...	1	2	3	4	5
28. Cuando oigo gente hablar en privado...	1	2	3	4	5
29. He recibido demasiado cambio de algún ...	1	2	3	4	5
30. Siempre declaro todo en ...	1	2	3	4	5
31. Cuando era joven, a veces ...	1	2	3	4	5
32. Nunca he tirado basura ...	1	2	3	4	5
33. En ocasiones manejo más rápido que el ...	1	2	3	4	5
34. Nunca leo libros o ...	1	2	3	4	5
35. He hecho cosas que no le digo ...	1	2	3	4	5
36. Nunca tomo cosas que no ...	1	2	3	4	5
37. Me he salido del trabajo o de la escuela por estar enfermo a ...	1	2	3	4	5
38. Nunca he dañado algún libro de biblioteca o mercancía de ...	1	2	3	4	5
39. Tengo algunos hábitos ...	1	2	3	4	5
40. No critico los asuntos de ...	1	2	3	4	5

ANEXO D

ESCALA DE DESEABILIDAD SOCIAL DE MARLOWE Y CROWNE

Marlowe & Crowne (DICO)		
Nombre: _____	Edad: _____	Género: <input type="checkbox"/> Masculino <input type="checkbox"/> Femenino
Ciudad de Residencia: _____	Fecha: ____/____/____ <small>Mes Día Año</small>	
Instrucciones: A continuación encontrara una serie de declaraciones en donde se requiere que consteste si es verdadera o falsa la afirmación, según corresponda con su persona. Marque el recuadro en la opción que refleje mejor su propio caso. NO HAY RESPUESTAS BUENAS NI MALAS. Por favor, conteste con toda franqueza, diciendo la verdad. Trate de no dejar ninguna pregunta sin contestar.		
	Verdadero (V)	Falso (F)
1. Antes de votar investigo detalladamente que tan ...	V	F
2. Nunca dudo salir de mi rutina para ayudar a...	V	F
3. En ocasiones me es difícil continuar con mi ...	V	F
4. Nunca he sentido un desagrado ...	V	F
5. En ocasiones he tenido dudas acerca de mis ...	V	F
6. En ocasiones me siento rencoroso cuando las ...	V	F
7. Siempre soy cuidadoso en cuanto a ...	V	F
8. Mis modales en la mesa en casa son igual ...	V	F
9. Si pudiera entrar al cine sin pagar y estar seguro ...	V	F
10. En algunas ocasiones he dejado de hacer algo por...	V	F
11. En ocasiones me gusta ...	V	F
12. Han habido ocasiones en las que he tenido ganas de rebelarme en contra...	V	F
13. No importa con quien esté hablando, siempre	V	F
14. Recuerdo "hacerme el enfermo"	V	F
15. Han habido ocasiones en las que me he ...	V	F
16. Siempre estoy dispuesto admitir ...	V	F
17. Siempre intento practicar ...	V	F
18. No encuentro particularmente difícil el llevarme con personas...	V	F
19. En ocasiones intento ajustar cuentas en vez de ...	V	F
20. Cuando no sé algo, no me ...	V	F
21. Siempre soy cortés, inclusive ...	V	F
22. En ocasiones he insistido arduamente para que las cosas ...	V	F
23. Han habido ocasiones en las que me he sentido con ganas ...	V	F
24. Nunca consideraría la posibilidad de que castiguen a alguien ...	V	F
25. Nunca he resentido el que se me pida que ...	V	F
26. Nunca me ha irritado el que la gente exprese ideas ...	V	F
27. Nunca hago un viaje largo sin haber revisado la ...	V	F
28. Han habido ocasiones en las que he estado muy celoso ...	V	F
29. Casi nunca he tenido la necesidad de ...	V	F
30. En ocasiones me irritan las personas....	V	F
31. Nunca he sentido que me ...	V	F
32. En ocasiones pienso que cuando le pasa algo malo a alguien ...	V	F
33. Nunca he dicho algo deliberadamente para lastimar ...	V	F

ANEXO E

Cuestionario alternativo para evaluación del instructor

Nombre del curso:	Nombre del instructor:	
Fecha de inicio:	Fecha de Término:	
<p>Su evaluación es muy importante para optimizar los servicios de Capacitación. Le agradeceremos llenar en forma objetiva y veras el presente cuestionario. Marque con una "X" la opción que mejor exprese su opinion del desempeño del instructor, usando los siguientes valores.</p>		
El instructor	NO	SI
1.- Domina el tema		
2.- Se expresó de forma clara		
3.- Despertó mi interés en los temas expuestos		
4.- Resolvió mis dudas		
5.- Relacionó los temas con mi trabajo		
6.- Me permitió intervenir		
7.- Sus conocimiento estan de acuerdo al objetivo del curso		
8.- Cumplió con el horario establecido		
9.- Cubrió en su totalidad con los temas programados		
10.- Cumplió satisfactoriamente el objetivo del curso		
Sugerencias y/o comentarios:		