

Informe general de la asignatura, mayo de 2016

PSICOLOGÍA (zona horaria 2)

(IB África, Europa y Oriente Medio e IB Asia-Pacífico)

Para proteger la integridad de los exámenes, cada vez es más frecuente el uso de variantes de los cuestionarios de examen según la zona horaria donde se realicen. Al recurrir a variantes del mismo examen, los alumnos ubicados en una parte del mundo no responderán siempre al mismo cuestionario de examen que los alumnos ubicados en otras partes del mundo. Se sigue un proceso muy riguroso para garantizar que los distintos cuestionarios de examen sean comparables en lo que respecta a su dificultad y a la cobertura del programa de estudios, y se toman las medidas pertinentes para garantizar que se apliquen las mismas normas de calificación a todos los exámenes escritos de los alumnos, independientemente de cuál haya sido la versión del examen a la que hayan respondido. En la convocatoria de exámenes de mayo de 2016, el IB preparó versiones distintas de la prueba 1 de Psicología para las diferentes zonas horarias.

Límites de calificación de la asignatura

Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-9	10-19	20-30	31-42	43-53	54-65	66-100

Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-9	10-22	23-32	33-43	44-54	55-65	66-100

Evaluación interna del Nivel Superior

Límites de calificación del componente

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-2	3-5	6-9	10-13	14-17	18-21	22-28

Ámbito y adecuación del trabajo entregado

En conjunto, el ámbito y la adecuación del trabajo entregado fueron de un buen nivel esta convocatoria, con una comprensión clara de los requisitos de la evaluación interna. La mayoría de los alumnos estaba al tanto de los problemas éticos e incluyó una copia del consentimiento informado y de las instrucciones para la reunión informativa y el informe posterior (*debriefing*) en los apéndices. Además, los alumnos conocían los requisitos de un diseño experimental, a pesar de que en esta convocatoria hubo un aumento considerable de la cantidad de experimentos casi/naturales, que no resultan aceptables ya que no se da una auténtica manipulación de una variable independiente.

La mayoría de los trabajos se basó en estudios de la psicología cognitiva y esto parece proporcionar buenos informes en este nivel de estudios. Los experimentos más utilizados fueron, como de costumbre, la memoria reconstructiva y los experimentos relativos a la teoría de los esquemas mentales, los estudios relativos a la duración de la memoria a corto plazo y la visualización frente a la repetición.

Sin embargo, sigue habiendo algunos problemas con las evaluaciones internas, entre otros:

- Se manipularon tres o cuatro variables, cuando en la *Guía de Psicología* se recomienda un estudio experimental sencillo con solo dos condiciones.
- La predicción realizada en la hipótesis no se justificó adecuadamente, ni tampoco tuvo el respaldo de ninguna investigación de contexto.
- Por lo que se refiere a la estadística descriptiva, no se explicó su uso.
- Por lo que se refiere a la estadística inferencial, se identificaron tests, pero no estaban justificados. Además, muchos alumnos no incluyeron los datos brutos ni cálculos del test inferencial elegido.
- Las discusiones eran superficiales y se incluía una discusión limitada de los resultados de la evaluación interna a la luz de la investigación de contexto o no se incluía referencia alguna a la estadística. La identificación de las limitaciones del procedimiento no estaba relacionada con sugerencias de modificación.

Debe tenerse en cuenta que no se exige hacer una réplica exacta de un experimento. Una réplica parcial es adecuada, pero el experimento del alumno debe estar estrechamente relacionado con un experimento real.

Desempeño de los alumnos en relación con cada criterio

Criterio A: Introducción

En algunas evaluaciones internas, la investigación que se presentaba no estaba vinculada *explícitamente* con las hipótesis. Los alumnos siempre deben describir la investigación (teorías o estudios) e indicar cómo se vincula con las hipótesis. No se debe presentar una investigación contradictoria.

Se recomienda presentar tres estudios o teorías. Es importante que la investigación de contexto presentada conduzca de manera lógica hacia las hipótesis de investigación, y se explique y analice con suficiente profundidad. Esto permite la formulación de una hipótesis de investigación clara en la introducción que, a su vez, estimulará la discusión de los resultados a la luz de la investigación de contexto en la sección de discusión.

Parece que los alumnos tuvieron dificultades para escribir las hipótesis de manera clara. Las variables deben operativizarse, es decir, deben poder ser mensurables. Por ejemplo, al medir la memoria, es mejor escribir “número de palabras que se recuerdan” que “mejor memoria”. Además, la redacción de la hipótesis debe expresar claramente el resultado que se espera alcanzar.

Criterio B: Diseño

La mayoría de los alumnos indicó un diseño adecuado (medidas repetidas o diseño independiente), pero la elección del diseño no siempre se justificó debidamente (es decir, por qué se eligió un diseño concreto frente a otro). Esta información debe explicarse con claridad.

Varios alumnos tuvieron dificultades para la operativización de las variables dependiente e independiente (es decir, para hacerlas claramente mensurables).

Criterio C: Participantes

No siempre se identificó de manera adecuada la población objetivo, es decir, la población de la que se extrajo la muestra. Con frecuencia los alumnos confundieron la muestra con la población objetivo.

En conjunto, la mayoría de los alumnos incluyó las características pertinentes de los participantes, tales como la edad, el sexo, si eran daltónicos (para el efecto Stroop) o su nivel de dominio del inglés. No es necesario incluir detalles irrelevantes, como el estatus socioeconómico.

En gran medida, se eligió e identificó correctamente la técnica de muestreo, pero a menudo no se explicó el uso de la técnica.

Criterio D: Procedimiento

Los alumnos deben hacer referencia a todas las pautas éticas que siguieron. Además, es necesario que se haga referencia a todos los materiales indicados en los apéndices. Sin unas referencias adecuadas, no será posible replicar correctamente el experimento.

Los alumnos también deben dejar claro cómo se distribuyeron los grupos (si se utilizaron medidas independientes). También se recomienda la compensación cuando se utilice un diseño de medidas repetidas.

Criterio E: Resultados: estadística descriptiva

La mayoría de los alumnos incluyó un gráfico y una tabla, con suficiente rotulación y una descripción de los resultados; solo se exige una medida de tendencia central y una medida de dispersión. Como en convocatorias anteriores, únicamente las mejores evaluaciones internas explicaron el uso de la estadística descriptiva, es decir, por qué se eligió una medida concreta de tendencia central o de dispersión.

Criterio F: Resultados: estadística inferencial

La mayoría de los alumnos eligió un test apropiado y justificó el uso del test (en base al nivel de datos y al diseño). En ocasiones se eligieron pruebas t (lo cual es aceptable), pero a menudo este no era el test más adecuado según los aspectos concretos del experimento. Parece que en esta convocatoria hubo un aumento en el número de alumnos que incluyeron los datos brutos o los cálculos de los tests inferenciales. Es importante que se incluyan en los apéndices los datos brutos y todos los cálculos del test inferencial. Si los cálculos se llevan a cabo en línea, puede incluirse en los apéndices una captura de pantalla de los cálculos para documentarlos.

Varios alumnos no hicieron un enunciado de significación estadística o no se aceptó o se rechazó la hipótesis nula, lo cual es necesario para obtener la máxima puntuación.

Criterio G: Discusión

Como en convocatorias anteriores, parece que esta sección del informe presentó las mayores dificultades para los alumnos, ya que con frecuencia careció de desarrollo y análisis. Con frecuencia, solo se hizo referencia a los resultados de la evaluación interna, pero es necesario discutir también dichos resultados. Los alumnos deben hacer referencia siempre a *todas* las investigaciones que se presenten en la introducción y deben discutir las haciendo referencia a sus propios hallazgos.

Casi todos los alumnos presentaron limitaciones, pero a menudo de manera superficial, sin un análisis riguroso. Deben presentarse limitaciones que sean pertinentes para la investigación, no limitaciones de los experimentos de carácter general. No es necesario incluir los puntos fuertes del diseño y del procedimiento.

También es necesario que se incluya una conclusión.

Criterio H: Cita de fuentes

A menudo los alumnos no incluyeron las referencias de las investigaciones que mencionaban en la introducción. Además, los alumnos no utilizaron un método estándar para las citas, como el de la APA, o las referencias no estaban completas.

Criterio I: Formato del informe

Por lo general, el formato de los informes estaba bien hecho. Los apéndices estaban bien organizados y rotulados.

El resumen debe incluir un sumario del estudio, así como los resultados del estudio.

Recomendaciones para la enseñanza de futuros alumnos

- Se recomienda que los profesores ayuden a los alumnos a encontrar la investigación de contexto adecuada, es decir, el marco teórico y los estudios adecuados. Se recomienda buscar estudios experimentales relativamente sencillos de replicar.
- La investigación de contexto de la introducción se debe analizar con la suficiente profundidad como para que queden justificados claramente los objetivos de la investigación del alumno, y la hipótesis del experimento debe estar claramente vinculada con la investigación de contexto.
- Para la muestra, no es necesario que el número de participantes en el experimento sea mayor de 20 (diseño de medidas independientes) o 10 (diseño de medidas repetidas).
- Resultaría útil proporcionar a los alumnos experimentos anteriores para que los leyeran y se familiarizaran con los aspectos de la investigación experimental. A algunos alumnos les beneficiaría llevar a cabo una “evaluación interna piloto” para familiarizarse con el formato y el procedimiento de un diseño experimental.
- Se recomienda encarecidamente que tanto el profesor como los alumnos se lean con atención la *Guía de Psicología* para asegurarse de que cumplen todos los requisitos.
- Debe enseñarse a los alumnos a hacer bien las referencias a la investigación, ya que con frecuencia la cita de las fuentes fue incompleta o no se presentó con la debida coherencia.
- Se recomienda en general que los alumnos se familiaricen con los estándares científicos, y debe fomentarse la lectura de investigaciones de contexto adecuadas. Se recomienda que se prepare a los alumnos para un uso crítico de los recursos presentes en Internet. Muchos alumnos solo utilizaron fuentes de Internet de carácter no especializado como bibliografía de apoyo.

Evaluación interna del Nivel Medio

Límites de calificación del componente

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-2	3-5	6-8	9-10	11-12	13-14	15-20

Ámbito y adecuación del trabajo entregado

Los trabajos presentados fueron muy variados en cuanto a calidad y variedad. Los trabajos que versaban sobre el nivel de análisis cognitivo fueron la opción más popular, y la mayoría de los trabajos entregados fue una réplica de estudios sobre temas de cognición y memoria en psicología. Resultó satisfactorio constatar que algunos alumnos replicaron investigaciones recientes. En la mayoría de los casos, los trabajos entregados resultaron adecuados para los alumnos del Programa del Diploma que estudian Psicología de Nivel Medio; además, respetaban las pautas éticas. Sin embargo, los moderadores se percataron de que algunos alumnos replicaron estudios demasiado complicados, lo que les hizo perder puntos en la sección de resultados, ya que no supieron analizar los resultados de manera eficaz y ajustarse a los criterios de evaluación.

La mayoría de los alumnos seleccionó estudios adecuados, supo describirlos bien y logró relacionar de algún modo sus propios resultados con los hallazgos del estudio original en la sección dedicada a la discusión.

Con frecuencia, los alumnos consiguieron algunos puntos en las secciones de introducción, diseño y participantes, si bien en ocasiones no incluyeron un objetivo o una justificación claros para el diseño o el método de muestreo, por lo que no lograron la puntuación máxima. En las secciones de procedimientos y resultados, los alumnos no lograron obtener parte de los puntos posibles porque no aportaron suficiente información precisa y pertinente. En algunos casos, los alumnos olvidaron ofrecer un relato verbal (descripción) de las estadísticas descriptivas que obtuvieron.

En muchos informes, los alumnos presentaron una sección de discusión redactada de manera superficial, por lo que perdieron muchos puntos debido a la falta de profundidad al discutir los hallazgos y las cuestiones metodológicas del estudio realizado.

Sorprendentemente, en esta convocatoria de mayo, hubo más ejemplos de informes que no cumplían los criterios para el trabajo experimental. Los alumnos tendieron a replicar estudios como el de Bartlett ("La guerra de los fantasmas"), o el de Brewer y Treyns sobre la memoria de los objetos en una sala, o llevaron a cabo un estudio de posición serial en el que midieron con qué frecuencia se memorizaba una palabra en relación con su posición en una lista. Aunque estos estudios representan un material de estudio significativo y pertinente en relación con algunos resultados del aprendizaje dentro del tronco común, no son auténticos

experimentos en los que la variable independiente sea manipulada explícitamente por la persona que realiza el experimento (el alumno). Los profesores deben informar a todos los alumnos de que es necesario asegurarse de que su estudio constituya un verdadero experimento. En la *Guía de Psicología* se define el estudio experimental sencillo de la siguiente manera: “El estudio consiste en la manipulación de una sola variable independiente y la medición del efecto que esta tiene sobre una variable dependiente, a la vez que controla otras variables”.

Desempeño de los alumnos en relación con cada criterio

Hubo numerosas muestras excelentes que demostraron un alto nivel de conocimiento de la metodología de investigación. En el extremo inferior, resultó evidente que algunos alumnos no tenían suficientes habilidades de gestión del tiempo, si bien se les habían enseñado adecuadamente, y por ello no dedicaron suficiente tiempo y esfuerzo a la realización de una buena investigación ni a la redacción de un informe detallado.

Criterio A: Introducción

En muchos casos, la introducción del informe estaba bien escrita, y la mayoría de los alumnos identificó y explicó claramente el estudio para la replicación parcial, y también logró presentar un objetivo expresado con claridad. A pesar de todo, algunos alumnos no lograron hacer lo que se requería y en su lugar incluyeron explicaciones y estudios adicionales, por lo que no identificaron de manera clara el estudio que estaban replicando. No es necesario proporcionar una revisión de varios estudios en la sección de introducción. Otro problema que los moderadores percibieron en ocasiones es que los alumnos siguieron los requisitos del Nivel Superior y redactaron hipótesis, por lo que se redujo el cómputo de palabras disponible para describir el experimento original o se complicó la redacción de una discusión adecuada y minuciosa, al quedar un número limitado de palabras.

Además, en algunos casos, los objetivos no se enumeraron de manera clara y a veces ni siquiera se mencionaron. Los alumnos se deben asegurar de mencionar el objetivo con precisión y claridad. Tienen que establecer una conexión entre las dos condiciones de la variable independiente e indicar cómo medirán la variable dependiente.

Criterio B: Diseño

A pesar de que el enunciado de la variable independiente y la variable dependiente en general fue correcto, la operativización de estas supuso un problema en muchas ocasiones. En algunos informes era necesario indicar con más claridad la definición de la variable independiente y la variable dependiente ya que, con frecuencia, resultaron demasiado vagas e imprecisas. Otro error frecuente fue indicar únicamente una condición (normalmente la condición experimental).

Además, hubo alumnos de muchos centros que tuvieron problemas para identificar adecuadamente el diseño en sí. Por otro lado, fueron demasiados los alumnos que identificaron vagamente el diseño como “experimental o controlado en un laboratorio”. Parece que algunos alumnos no supieron distinguir entre el diseño y el método. Además, el uso del diseño no siempre se justificó eficazmente. En algunos casos, los alumnos ofrecieron

justificaciones incorrectas o no ofrecieron ninguna justificación para su elección del diseño experimental.

En general, hubo una mejora notable en la identificación y la discusión de consideraciones éticas (consentimiento informado, *debriefing*, etc.). En pocas ocasiones, los alumnos no proporcionaron formularios de consentimiento de los padres cuando los participantes eran menores de 16 años.

Criterio C: Participantes

En muchos casos, los alumnos ofrecieron buenas descripciones de características pertinentes de los participantes, incluidas las características de la población objetivo adecuadas y la identificación de su técnica de muestreo. Sin embargo, muchos alumnos no justificaron su uso de esta técnica de muestreo y por ello no obtuvieron la puntuación máxima. El término "aleatorio" parece seguir siendo una fuente de confusión que se refleja en la descripción de la selección de participantes y la asignación de condiciones. Muchos alumnos tendieron a utilizarlo como explicación de un procedimiento de muestreo que no tiene sesgo. El muestreo aleatorio es mucho más que eso y significa que todos los participantes tienen las mismas posibilidades de ser seleccionados. En ocasiones, los alumnos se esforzaron por tener un número alto de participantes; sin embargo, no es necesario, basta con que sus estudios incluyan entre 15 y 20 participantes. Menos de 15 resulta arriesgado, ya que podría haber un problema provocado por resultados no esperados. Por otra parte, no se recomiendan las muestras de tamaño considerable, ya que no resultan prácticas.

Aunque buena parte de los alumnos fue capaz de dar alguna característica de la muestra, en algunos casos no se logró identificar características pertinentes de los participantes.

Criterio D: Procedimiento

En la mayoría de los casos, los procedimientos fueron pertinentes y se describieron con claridad, pero en algunos casos los materiales a los que se hizo referencia no estaban incluidos en los apéndices (por ejemplo, instrucciones estandarizadas, tests, cuestionarios), lo que afectó a la posibilidad de repetir el procedimiento. Si bien en general esta sección del informe estaba bien escrita, aún se puede mejorar la redacción. Además, fueron pocas las ocasiones en que hubo un *debriefing* completo y detallado.

Criterio E: Resultados

La sección de resultados tendió a ser difícil para muchos alumnos, y los moderadores raramente pudieron otorgar la puntuación máxima. Los profesores deben instruir claramente a los alumnos sobre lo que es obligatorio y necesario en esta sección.

Hubo una variedad muy pronunciada en las secciones de resultados, algo relacionado con la naturaleza del estudio que se replicaba. Los estudios sencillos permitieron que los alumnos cumplieran los criterios de evaluación, pero los estudios complicados provocaron problemas a los alumnos a la hora de presentar sus resultados. Con demasiada frecuencia, no se explicó ni se justificó el uso de medidas de tendencia central y dispersión. Además, algunos alumnos no describieron de manera eficaz los resultados por escrito. Otro problema frecuente fue que

algunos alumnos incluyeron en esta sección datos brutos o gráficos que mostraban los resultados de cada participante.

En muchos informes, los gráficos no estaban rotulados con suficiente claridad como para poder reconocerlos, ya que los rótulos resultaban inexactos o incompletos. Los alumnos con peores resultados eligieron el tipo de gráfico erróneo (histogramas o gráficos circulares para mostrar diferencias entre grupos independientes). Además, varios alumnos presentaron sus resultados con poca claridad, ya que no incluyeron porcentajes, medidas de tendencia central ni dispersión. Algunos alumnos proporcionaron distintos gráficos en la sección de resultados y presentaron los datos de distintas maneras, aunque muchas veces no reflejaron el objetivo del estudio. Además, en ocasiones se aplicó la estadística de manera incorrecta. Por ejemplo, cuando se utilizaron niveles ordinales de medición, algunos alumnos buscaron la puntuación media a pesar de que no resulta una medida de tendencia central adecuada para datos ordinales.

En la sección de resultados, los alumnos se deben asegurar de proporcionar los títulos de las tablas y las figuras, además de una descripción suficiente de lo que reflejan. Es importante que los alumnos nombren específicamente sus medidas de tendencia central: ¿reflejan la media, la mediana y la moda? Los alumnos deben describir lo que reflejan estas puntuaciones distintas para los grupos experimentales y de control, y es muy importante describir lo que implica la desviación típica o el rango.

Criterio F: Discusión

Como es habitual, la calidad de la sección de discusión fue variada. En esta convocatoria de mayo hubo algunas mejoras, ya que las discusiones tendieron a seguir con más claridad los criterios para esta sección. Hubo muchos más alumnos que relacionaron la discusión de los puntos débiles con el tipo de diseño elegido. Además, los alumnos tendieron a incluir las conclusiones en la sección de discusión en vez de al final del informe.

Desafortunadamente, los alumnos que no habían descrito claramente el estudio que se había replicado en la introducción también tendieron a tener dificultades con la sección de discusión.

Al igual que en años anteriores, los alumnos con peores resultados tuvieron problemas con la sección de discusión, sobre todo para relacionar sus resultados con el estudio replicado y evaluar sus propios diseños y procedimientos. La mayoría de las discusiones giró en torno a cuestiones menos pertinentes, como el método de muestreo y el comportamiento (bueno o malo) de los participantes. Además, algunos alumnos destacaron los puntos fuertes de su estudio aunque no era necesario. Un problema frecuente de algunos informes fue que los alumnos no lograron abordar con claridad las limitaciones; además, a menudo se omitieron las sugerencias para investigaciones adicionales, o apenas se sugirieron.

Criterio G: Presentación

En general, los informes respetaron el límite máximo de palabras, aunque en algunas ocasiones los alumnos no anotaron el cómputo de palabras. En la mayoría de los casos, los informes utilizaron el formato requerido, y se proporcionaron referencias. Sin embargo, muchas veces no se proporcionaron los datos completos de la publicación de manera coherente o

completa; se debe instar a los alumnos a que utilicen un sistema de referencia estándar. Los resúmenes variaron en calidad, oscilando desde aquellos que eran claros y concisos hasta otros que incluían muy poca información concreta y relacionada con la réplica de un estudio original. En ocasiones, pareció que algunos alumnos finalizaron sus informes apresuradamente, por lo que omitieron algunos elementos en los apéndices (por ejemplo, los materiales utilizados, las instrucciones estandarizadas, el formulario de consentimiento o los cálculos realizados). Hubo demasiados alumnos que no rotularon bien los apéndices y no se refirieron a ellos o no incluyeron los documentos necesarios en los apéndices. Los informes se mantuvieron por lo general dentro del límite de palabras.

Recomendaciones para la enseñanza de futuros alumnos

- Los profesores deben ser claros sobre cuáles son los requisitos básicos de la evaluación interna en lo que se refiere a los temas o experimentos que no resulta adecuado replicar por cuestiones éticas; de ese modo, podrán orientar a los alumnos para que elijan de manera adecuada.
- Además, al elegir un experimento, es importante que los alumnos sean capaces de entender todos los aspectos del estudio y centrar el objetivo en lo que es factible dentro de los criterios de evaluación.
- Contar con más información sobre las ventajas y los inconvenientes de utilizar distintos diseños experimentales y técnicas de muestreo ayudará a los alumnos a justificar su uso en los informes.
- Un estudio experimental para el Nivel Medio no debe tener más de una variable independiente y una variable dependiente.
- En la introducción, debe decirse a los alumnos que escriban únicamente acerca del estudio que se ha replicado y que luego indiquen de forma clara su propio objetivo al final.
- La sección de resultados debe proporcionar claramente estadísticas descriptivas relacionadas con el objetivo del estudio.
- En la sección de resultados, debe decirse a los alumnos que incluyan solamente la tabla con las estadísticas descriptivas y un gráfico, y que justifiquen o expliquen su elección de estadísticas descriptivas. Los alumnos también necesitan orientación más específica sobre cómo describir los resultados por escrito.
- Debe recomendarse a los alumnos que comprueben todos los cálculos y los incluyan en los apéndices.
- Se debe hacer más hincapié en la importancia de una discusión bien equilibrada que establezca conexiones explícitas entre la metodología y los resultados de su estudio.
- En la sección de discusión, los alumnos deben hacer referencia a sus medidas de tendencia central y dispersión, así como comparar significativamente sus propios hallazgos con el estudio que se ha replicado.
- Los alumnos necesitan orientación más específica y ejemplos sobre cómo centrar sus discusiones en el diseño y el procedimiento y no en el muestreo y el comportamiento de los participantes. Además, necesitan más instrucción sobre cómo identificar limitaciones significativas y pertinentes y establecer conexiones con las modificaciones para abordar esas limitaciones específicas.
- Hace falta más orientación acerca del formato esperado para la evaluación interna (por ejemplo, saber dónde deberían comentarse las consideraciones éticas, dónde

deberían presentarse los datos brutos y dónde deben ir las instrucciones estandarizadas).

- Se debe alentar a los alumnos a que revisen sus informes antes de entregarlos.

Prueba 1 del Nivel Superior y del Nivel Medio

Límites de calificación del componente del Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-4	5-9	10-14	15-19	20-23	24-28	29-46

Límites de calificación del componente del Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-4	5-9	10-14	15-19	20-23	24-28	29-46

Áreas del programa y del examen que resultaron difíciles a los alumnos

Muchos alumnos tuvieron problemas para comprender qué exigía la pregunta; muchos aplicaron una única estrategia para todas las preguntas, independientemente de cuál fuera el término de instrucción. Esto incluía la redacción de los objetivos del nivel de análisis, la redacción de una historia del nivel de análisis y, en el caso de las preguntas de respuesta breve, la evaluación de la investigación. Muchas veces las introducciones largas en las preguntas de respuesta corta resultaron irrelevantes para lo que exigía la pregunta.

Muchos alumnos ofrecieron una evaluación de la investigación superficialmente pertinente o poco desarrollada en las preguntas de respuesta larga. La evaluación no debe limitarse solamente a identificar los problemas; también debe explicarlos dentro del contexto del estudio que se discute y de la pregunta más amplia que se aborda. En muchas ocasiones, el pensamiento crítico fue un añadido a un ensayo descriptivo. El pensamiento crítico no se debe limitar a una evaluación de la metodología y la ética de un estudio, sino que debe incluir una síntesis de los conceptos y estudios, el análisis de la teoría y los hallazgos, y la aplicación en otros estudios.

En muchas ocasiones, los alumnos demostraron malentendidos básicos acerca de la investigación psicológica. Muchos alumnos argumentaron que los psicólogos han demostrado que las teorías son ciertas, que no se puede generalizar la investigación y que la investigación realizada en laboratorios no puede predecir lo que sucede fuera de ellos.

En varias ocasiones, los alumnos redactaron sus ensayos en un solo párrafo, sin hacer una transición clara entre ideas o sin desarrollar un argumento.

Áreas del programa y del examen en las que los alumnos demostraron estar bien preparados

El conocimiento general de los tres niveles de análisis resultó satisfactorio, y la mayoría de los alumnos supo identificar conceptos adecuados y explicaciones teóricas, así como proporcionar ejemplos de investigación adecuados.

Puntos fuertes y débiles de los alumnos al abordar cada una de las preguntas

Sección A

Nivel de análisis biológico

La mayoría de los alumnos supo identificar una hormona. Fueron muy pocos los que identificaron un neurotransmisor en vez de una hormona. Muchos alumnos demostraron un buen conocimiento de la función de una hormona y de sus posibles efectos. Con frecuencia se utilizaron investigaciones pertinentes de manera eficaz para respaldar la respuesta.

Muchos alumnos se limitaron a describir estudios sin abordar la función de la hormona. Muchos alumnos no conocían la función del cortisol en la reacción ante el estrés, es decir, aumentar el nivel de azúcar en sangre, suprimir el sistema inmunitario y ayudar en el metabolismo de las grasas, las proteínas y los carbohidratos. Además, el cortisol ayuda a que el cuerpo vuelva a un estado de homeostasis tras una situación estresante. Muchos alumnos simplemente escribieron “mantener la homeostasis”. Además, en varias respuestas sobre la adrenalina los alumnos se centraron en el estudio de Schachter y Singer, pero no abordaron la función de la hormona. La pregunta se centraba en el comportamiento humano, pero muchos alumnos únicamente utilizaron ejemplos con animales sin vincularlos explícitamente con el comportamiento humano.

Nivel de análisis cognitivo

Hubo muy buenas respuestas en las que los alumnos demostraron que comprendían con claridad tanto la teoría de los esquemas como un estudio pertinente. Sin embargo, los alumnos no siempre tuvieron clara la naturaleza de los procedimientos o solamente tenían una comprensión general de los resultados.

Nivel de análisis sociocultural

Esta pregunta resultó difícil para algunos alumnos. Muchas respuestas se centraron en el efecto de los estereotipos en vez de en la manera en que se forman. Además, algunos alumnos tuvieron dificultades para establecer la conexión entre la teoría de la identidad social y la formación de estereotipos, y se limitaron a describir la teoría con gran detalle, sin comprender cómo puede llevar a los estereotipos.

Sección B

Nivel de análisis biológico

A muchos alumnos esta pregunta les resultó difícil. Un gran número de respuestas se limitó a describir y evaluar la investigación biológica en vez de describir y discutir las consideraciones éticas. Con frecuencia, las cuestiones éticas solamente se identificaron, sin adentrarse en una discusión clara sobre la naturaleza de las consideraciones, por qué estas resultan significativas en la investigación biológica o el proceso por el cual se toman las decisiones.

Muchos alumnos eligieron investigaciones muy antiguas, por ejemplo, la de Phineas Gage, Broca o Bouchard, lo que dio como resultado discusiones superficiales sobre el consentimiento, la confidencialidad y el *debriefing*. Además, parece que muchos alumnos no entendían correctamente cómo se obtiene el consentimiento informado en los estudios de caso de HM y Bouchard. En las mejores respuestas se discutió sobre por qué se utilizan placebos en los experimentos, por qué se debe mantener la confidencialidad en los estudios de casos o sobre la aplicación de la tres R (reemplazar, reducir y refinar) en la investigación animal.

Nivel de análisis cognitivo

Hubo algunas respuestas muy buenas a esta pregunta, en las que los alumnos demostraron entender con claridad cómo la biología y la cognición interactúan en la emoción. Algunos alumnos no se centraron en la naturaleza de la emoción, sino que abordaron la memoria “flashbulb”, algo que no resultaba adecuado para esta pregunta.

Los alumnos mejor preparados se centraron en la manera en que los psicólogos miden y comprenden la interacción de los factores biológicos y cognitivos. El pensamiento crítico centrado en las consideraciones éticas apenas tuvo una pertinencia marginal para esta pregunta.

Nivel de análisis sociocultural

Hubo algunas respuestas muy buenas que demostraron una amplitud de conocimientos y un pensamiento crítico bien desarrollado. Sin embargo, en muchas ocasiones las respuestas resultaron demasiado descriptivas y el pensamiento crítico se añadió a última hora. En muchos ensayos se aplicaron repetidamente estrategias de evaluación superficial, tales como “al tratarse de un experimento de laboratorio, carecía de validez ecológica”.

Los alumnos también tuvieron dificultades para escribir definiciones claras de los términos. Por ejemplo, con frecuencia los factores disposicionales se describieron como “factores que forman parte de nuestra disposición”. Para poder demostrar una comprensión clara, los alumnos no deberían utilizar el término para definir el propio término.

Recomendaciones y orientación para la enseñanza de futuros alumnos

Los profesores deben centrarse en el desarrollo de las habilidades de redacción de ensayos. Es importante que los alumnos se den cuenta de que no hay una estrategia única que les

permita responder a todas las preguntas. Los profesores deben comentar los trabajos con los alumnos, lo que incluye ayudarlos a entender lo que es innecesario o irrelevante en sus respuestas. Los profesores deben revisar claramente los términos de instrucción para que los alumnos se centren correctamente en los requisitos de las preguntas.

Al describir los estudios, los alumnos deberían ser capaces de resumir claramente el objetivo, el procedimiento y los hallazgos. Esto resulta necesario para el lector sin conocimientos del tema. Muchos alumnos redactaron sus respuestas partiendo de que el lector (es decir, el examinador) conocería el estudio. Por otro lado, muchos alumnos demostraron una comprensión superficial de los estudios y comprendían con limitaciones el procedimiento y cómo los hallazgos están relacionados con la pregunta abordada.

Los alumnos deben practicar la definición de términos y la descripción de teorías, y no deberían utilizar el término para definir el propio término.

Es importante elegir cuidadosamente la investigación que se enseña. El objetivo del programa es que los alumnos alcancen una buena comprensión de la psicología moderna, no de la historia de la psicología. Por tanto, se recomienda encarecidamente que no se utilicen las investigaciones del siglo XIX como centro de lo que se enseña.

Finalmente, se debe practicar más el pensamiento crítico, y también es necesario desarrollarlo más en las evaluaciones. El pensamiento crítico debe constituir una parte significativa de la respuesta y no la frase final de un párrafo. Para ayudar a desarrollar esta habilidad resultaría útil dar a los alumnos un estudio y pedirles que redacten un párrafo en el que evalúen la metodología y también los vínculos con la pregunta abordada.

Prueba 2 del Nivel Superior y del Nivel Medio

Límites de calificación del componente del Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-6	7-12	13-16	17-21	22-26	27-31	32-44

Límites de calificación del componente del Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-2	3-5	6-6	7-9	10-12	13-15	16-22

Áreas del programa y del examen que resultaron difíciles a los alumnos

A menudo los alumnos tuvieron dificultades para comprender los términos de instrucción, y esto complicó la tarea de abordar las preguntas con un enfoque adecuado. Esto supuso un problema, sobre todo para las preguntas en las que los alumnos tenían que “discutir” y “comparar y contrastar”. En muchos casos, los alumnos ofrecieron descripciones paralelas de tratamientos, teorías o modelos opuestos en vez de auténticas comparaciones y contrastes. Los indicios de pensamiento crítico en la mayoría de respuestas fueron limitados o débiles, y las explicaciones a menudo carecían de profundidad en la mayoría de los trabajos. Además, en las preguntas con el término de instrucción “evaluar”, a los alumnos les pareció difícil dirigir su evaluación hacia el concepto en vez de realizar la evaluación del estudio empírico al establecer una conexión entre el estudio empírico y el concepto evaluado.

Áreas del programa y del examen en las que los alumnos demostraron estar bien preparados

En general, los alumnos mantuvieron una buena estructura del ensayo (introducción, cuerpo principal y conclusión) durante toda la respuesta, incluso cuando el centro de la respuesta iba variando. La mayoría de las respuestas reveló una buena comprensión y conocimiento de los conceptos. Las respuestas de las dos opciones más populares (psicología del comportamiento anormal y psicología de las relaciones humanas) estaban muy bien organizadas, con un argumento equilibrado. Casi todas las respuestas ofrecieron claramente evidencia empírica. Las preguntas 1, 7 y 11 demostraron ser difíciles, y los alumnos que abordaron estas preguntas demostraron un buen conocimiento y comprensión de los estudios o las teorías psicológicas pertinentes.

Puntos fuertes y débiles de los alumnos al abordar cada una de las preguntas

Psicología del comportamiento anormal

Pregunta 1

Esta fue una pregunta muy popular; en muchas ocasiones las respuestas reflejaron un buen conocimiento de los enfoques del tratamiento. Los alumnos eligieron una variedad de distintos enfoques. Los más populares fueron los enfoques médicos y los enfoques cognitivos del tratamiento. Otros enfoques elegidos incluyeron: la desensibilización sistemática, la terapia de grupo, la terapia ecléctica, la terapia familiar y la terapia cognitiva basada en la conciencia plena (MBCT, por sus siglas en inglés), entre otros. Las mejores respuestas ofrecieron dos enfoques del tratamiento de un trastorno; además, hubo una serie de puntos que abordaron al comparar y contrastar. Algunos alumnos eligieron dos enfoques amplios del tratamiento. Si bien este enfoque resulta perfectamente aceptable, a menudo dio como resultado varias respuestas que reflejan un gran conocimiento detallado, pero sin demasiados indicios de comprensión o pensamiento crítico. En la mayoría de los casos, las dos primeras páginas de las respuestas incluían largas descripciones generales y evaluaciones de los enfoques del tratamiento donde solo el párrafo de conclusión final centraba la respuesta con claridad y de manera directa en la pregunta, al presentar las similitudes y diferencias entre los enfoques. Además, el principal problema que hubo en las respuestas estaba relacionado con el hecho de no abordar explícitamente el término de instrucción.

Otro problema se debió a que algunos alumnos dieron una descripción y evaluación generales y vagas de dos enfoques del tratamiento distintos, sin relacionarlas con un trastorno concreto.

En las preguntas de mayor calidad, se ofrecieron indicios de pensamiento crítico, ya que se comparó y se contrastó la eficacia de los dos enfoques, se discutió la adecuación para distintos contextos culturales y se abordaron consideraciones éticas.

Pregunta 2

Esta pregunta también resultó popular, y la calidad de las respuestas fue muy variada, con respuestas que iban desde las muy básicas y simplistas a las que discutieron rigurosamente varias consideraciones éticas en profundidad. Con frecuencia, las respuestas incluyeron los estudios siguientes: el estudio de Rosenhan sobre “estar cuerdo en lugares dementes”, el estudio de Cooper sobre diferencias culturales en el diagnóstico o el estudio de Broverman et al. sobre sesgos de género. Normalmente, las preguntas de mayor calidad se centraron únicamente en distintas consideraciones éticas que los alumnos luego discutieron en profundidad. Entre las consideraciones éticas más populares que se discutieron se incluyen: la profecía autocumplida, los efectos de la identificación, la posibilidad de la estigmatización, el sobrediagnóstico de determinados trastornos relacionados con el género y la cultura, y los sesgos biológicos. Desafortunadamente, algunos alumnos no entendieron bien la pregunta y discutieron las consideraciones éticas en los estudios de investigación en vez de en el diagnóstico.

Pregunta 3

La calidad de las respuestas fue variada, ya que algunos alumnos proporcionaron respuestas simples y basadas en el sentido común, mientras que otros se adentraron en discusiones sobre las variaciones de género en la prevalencia de distintos trastornos. La mayoría de los alumnos decidió centrarse en más de un trastorno. Los trastornos más populares fueron: trastorno por estrés postraumático (TEPT), trastornos de la alimentación, depresión y, en ocasiones, las fobias. Las respuestas de mayor calidad hicieron referencias claras a estudios pertinentes, incluidos los estudios relacionados con los modelos de vulnerabilidad y los estresores vitales (Brown y Harris, 1978) y el estudio sobre las diferencias en los estilos cognitivos (Nolen-Hoeksema, 1994).

Un punto débil frecuente fue que el término de instrucción “discutir” no se enfocó bien, lo que ocasionó que las respuestas resultaran demasiado descriptivas.

Psicología del desarrollo

Pregunta 4

Esta pregunta fue la opción más elegida en la parte de psicología del desarrollo. La mayoría de los alumnos eligió comparar y contrastar el modelo de asimilación y acomodación del desarrollo cognitivo de Piaget y el enfoque contextual del desarrollo cognitivo de Vygotsky. Estas respuestas tendieron a reflejar un buen conocimiento de los conceptos principales proporcionados por estos autores. Sin embargo, en muchos casos la obra de Piaget se describió con gran detalle, mientras que el trabajo de Vygotsky no estuvo bien representado. En algunos casos, los alumnos eligieron erróneamente teorías que eran pertinentes para la psicología del desarrollo, pero que no lo eran para el desarrollo cognitivo. Algunos ejemplos de elecciones erróneas incluyen: la teoría del apego de Bowlby, las etapas de la vida de Erikson y la teoría del aprendizaje social. En la mayoría de las respuestas hubo indicios de que los alumnos tenían un buen conocimiento y comprensión de una teoría, así como algunos indicios de que se conocía la segunda teoría. Sin embargo, muchos alumnos no supieron comparar y contrastar con claridad las dos teorías y, en su lugar, proporcionaron fundamentación empírica y evaluaron las teorías. Desafortunadamente, en estos casos, las puntuaciones para el criterio B se encontraron en la banda inferior, ya que los indicios de pensamiento crítico no resultaron pertinentes para la pregunta en cuestión.

Las respuestas de gran calidad establecieron puntos de comparación y contraste pertinentes relacionados con las suposiciones de las teorías, las consideraciones metodológicas de la investigación empírica, y la productividad de las teorías en relación con la aplicación y generación de investigación psicológica.

Pregunta 5

Esta fue la segunda pregunta más popular en esta opción. En general, esta pregunta se abordó con indicios limitados de conocimiento y comprensión. En la mayoría de las respuestas se hizo referencia a las siguientes teorías: la teoría de los esquemas de género que destaca el papel clave de los procesos cognitivos en el desarrollo de los roles de género, la teoría del aprendizaje social que pone de relieve la importancia del entorno social y hace hincapié en la

potencia de los procesos de observación y modelización, y la teoría de la diferenciación psicosexual basada en la suposición de que los roles de género están relacionados con el sexo genético determinado por los cromosomas. Algunas respuestas menos satisfactorias intentaron abordar la pregunta refiriéndose a la teoría evolutiva o a la teoría psicodinámica.

A continuación se presentan los estudios más populares seleccionados y descritos por los alumnos:

- Estudio de Martin y Halvorson (1983) que muestra el rol de los esquemas de género en los roles de género
- Estudio de Martin y Halvorson (1978) que muestra la influencia de los padres en los roles de género
- Estudio de Mead (1935) que muestra que los roles de género dependen de la sociedad
- Estudio de Money y Ehrhardt (1972) que sostiene que los niños al nacer tienen un género neutro

Muchos alumnos tuvieron problemas para proporcionar indicios claros de pensamiento crítico relacionados con la pregunta. Sin embargo, en los casos en los que se incluyó una discusión clara, los alumnos se refirieron a lo siguiente:

- Indicios complementarios y contradictorios
- El papel de los factores socioculturales, como los medios de comunicación, las influencias paternas o los estereotipos
- La interacción entre la naturaleza y la crianza
- Consideraciones metodológicas o éticas de los estudios complementarios

Pregunta 6

Muy pocos alumnos respondieron esta pregunta. La mayoría de las respuestas se centró en definiciones y discusiones sobre la resiliencia, y se prestó poca atención a las estrategias para desarrollarla. Parece que, en la mayoría de los casos, los alumnos peor preparados eligieron esta pregunta. También parece que muchos alumnos no habían entendido el concepto de resiliencia. Con frecuencia las respuestas de menor calidad incluyeron la descripción del estudio de caso de Koluchova acerca de los gemelos idénticos checoslovacos, el estudio de Genie e incluso a veces el estudio de caso de David Reimer a cargo de Money sin relacionar estos estudios con el tema de la resiliencia. Parece que esta pregunta puede haber alentado a los alumnos a ofrecer respuestas diferentes sobre el efecto del trauma y la privación en el desarrollo posterior. Las mejores respuestas hicieron referencia a los siguientes temas:

- Programas sociales para los jóvenes, como “Head Start” o “the Big Brothers Big Sisters Programme” (Tierney et al., 1985)
- Programas sobre la educación parental (Sanders et al., 2002)
- Programas que desarrollan habilidades para proteger y promover el bienestar (por ejemplo, la terapia cognitivo-conductual y la capacitación sobre habilidades sociales)
- Capacitación para inocular el estrés

En la mayoría de los casos, los alumnos intentaron abordar la evaluación de manera limitada, y casi siempre incluyeron referencias a las consideraciones metodológicas, culturales y de género, así como a cuestiones de edad o madurez de la persona.

Psicología de la salud

Pregunta 7

Esta fue una elección muy popular en esta opción. Muchas respuestas reflejaron un buen conocimiento y comprensión de los estudios y teorías pertinentes relacionados con los aspectos fisiológicos y psicológicos del estrés. Los aspectos fisiológicos del estrés más abordados fueron los siguientes:

- El papel del cerebro en el desarrollo del estrés y los mecanismos que existen en el cerebro que aspiran a minimizar el estrés (Hegel et al., 1989)
- Las respuestas suprarrenales a los estresores ambientales
- La función de la reducción de cortisol en el trastorno por estrés postraumático (TEPT)
- La conexión entre el estrés y el sistema inmunológico

Los aspectos psicológicos del estrés más abordados fueron los siguientes:

- La manera en que una persona valora una situación, es decir, la valoración cognitiva
- Estilo atribucional, tanto positivo como negativo
- Amenazas percibidas hacia el “yo social”
- Papel de la personalidad en la gestión del estrés

Los estudios y las teorías de investigación también estuvieron presentes a menudo, e incluyeron las siguientes elecciones:

- Teoría de lucha o huida de Cannon (1914)
- Modelo del síndrome de adaptación general de Selye (1956)
- Experimento natural de Kiecolt-Glaser et al. (1984) para investigar si el estrés provocado por un examen importante tuvo un efecto en el funcionamiento del sistema inmunológico del cuerpo

Las mejores respuestas tendieron a tener en cuenta una pequeña cantidad de aspectos fisiológicos y psicológicos del estrés, por lo que demostraron profundidad de conocimiento.

Un punto débil frecuente en las respuestas fue que el término de instrucción “discutir” no se enfocó bien, lo que ocasionó que las respuestas resultaran demasiado descriptivas. En algunos casos (raros) los alumnos solamente discutieron los aspectos fisiológicos y psicológicos del estrés.

Pregunta 8

Esta pregunta no fue muy popular y principalmente atrajo la atención de los alumnos menos preparados. Con mucha frecuencia, las respuestas resultaron demasiado generales y vagas, y presentaron indicios mínimos de pensamiento crítico.

Los factores socioculturales abordados con más frecuencia fueron la posición socioeconómica, la educación, las normas culturales, las normas socioculturales y la influencia de los medios de comunicación.

Algunos alumnos también se refirieron a los factores biológicos o cognitivos para abordar el término de instrucción “en qué medida”. Este enfoque normalmente resultó adecuado, y estas respuestas tendieron a obtener una puntuación por encima de la media.

Pregunta 9

Parece que los alumnos ignoraron con frecuencia el término de instrucción “evaluar”, ya que solo proporcionaron descripciones exhaustivas de tratamientos pertinentes.

En la mayoría de los casos, los alumnos hicieron referencia al tratamiento sustitutivo con nicotina, el tratamiento contra las drogas, la reducción del estrés basada en la atención plena, el tratamiento ecléctico o las terapias en grupo.

En algunos casos, los alumnos parecieron ignorar el número “dos” que aparecía en la pregunta, ya que, en su lugar, evaluaron uno o más de dos tratamientos; en ambos casos, las respuestas no lograron una puntuación alta, ya que no se centraron en lo que pedía específicamente la pregunta.

Psicología de las relaciones humanas

Pregunta 10

Esta fue una pregunta muy popular que no supuso problemas para la mayoría de los alumnos. Sin embargo, algunos alumnos se adentraron en una discusión sobre el comportamiento prosocial en vez de sobre las teorías del altruismo, acompañándose de una discusión demasiado larga sobre el caso de Kitty Genovese.

En la mayoría de las respuestas se eligió una de las siguientes teorías:

- Teoría del gen egoísta de Dawkin
- Teoría de la selección de parentesco
- Teoría del altruismo recíproco de Trivers
- Modelo del alivio del estado negativo de Cialdini
- Modelo de la empatía-altruismo de Batson

Muchas respuestas reflejaron un buen conocimiento y comprensión de una teoría que explicaba el altruismo en los humanos.

Sin embargo, un punto débil frecuente fue que los alumnos no lograron establecer una conexión entre la teoría, el concepto o el término y los estudios empíricos que proporcionaron. Por eso, la evaluación proporcionada no estaba orientada al concepto, sino a los estudios empíricos; en muchas ocasiones los requisitos de la pregunta no se abordaron directamente.

Pregunta 11

En muchas ocasiones los alumnos no abordaron con claridad el término de instrucción “analizar”; este término requiere que los alumnos destaquen (hagan hincapié en) los aspectos esenciales del papel que desempeña la cultura a la hora de formar y mantener relaciones.

Los alumnos abordaron esta pregunta haciendo referencia al papel del individualismo frente al colectivismo, la diferencia entre culturas continuas y culturas discontinuas, la idea de que en algunas culturas se valora más la castidad y las habilidades en el hogar para las mujeres, la teoría evolutiva que sugiere que hay patrones universales en la formación y el mantenimiento de las relaciones, y el papel de las normas culturales en la formación y el mantenimiento de las relaciones.

Entre los estudios frecuentemente incluidos en la respuesta estaban los siguientes:

- Estudio comparativo de Yelsma y Athappilly (1988) sobre los matrimonios concertados y los matrimonios por amor
- Estudio de Buss et al. (1990) sobre preferencias internacionales a la hora de seleccionar pareja (un estudio de 37 culturas)
- Estudio de Levine et al. (1995) sobre el papel del amor en el establecimiento del matrimonio
- Estudio intercultural de Buss (1994) sobre las relaciones
- Estudio de Ahmad y Reid (2008) sobre los estilos de comunicación en los matrimonios concertados

Normalmente se proporcionaron indicios de pensamiento crítico, ya que se incluyó una evaluación de la investigación pertinente y también se compararon y contrastaron las similitudes y las diferencias culturales en las relaciones, y se discutió la interacción entre los factores biológicos y los culturales.

Hubo varios alumnos que intentaron responder esta pregunta mediante la identificación y la descripción de estudios pertinentes; sin embargo, no fueron capaces de discutir de qué manera los hallazgos de los estudios resultaban pertinentes para la pregunta.

Pregunta 12

Esta fue una elección muy popular en esta opción y parece que atrajo más a los alumnos bien preparados.

Los alumnos abordaron claramente los efectos de la exposición a la violencia; para ello, se refirieron a las respuestas fisiológicas al estrés, a la reducción del cortisol que provoca fatiga crónica, a los efectos en la salud mental, al menor rendimiento en el colegio, a enfermedades psicosomáticas, al círculo de la violencia o a la delincuencia.

En las respuestas se incluyeron referencias a los siguientes estudios:

- Estudio de Shalev y Freedman (2005) sobre el trastorno por estrés postraumático tras ataques terroristas
- Estudio de Kumar et al. (2005) sobre el efecto de la violencia doméstica en la salud mental de las mujeres indias

- Estudio de Shuster et al. (2001) sobre las respuestas de estrés a la exposición a los ataques terroristas del 11 de septiembre de 2001
- Estudio de Hyman (1990) sobre la exposición a largo plazo y la depresión

Los alumnos reflejaron muy claramente el componente del conocimiento en sus respuestas, si bien parece que tuvieron más dificultades a la hora de proporcionar indicios claros de pensamiento crítico pertinente. Las respuestas de mayor calidad abordaron cuestiones como la dificultad para obtener pruebas empíricas, la dificultad para definir términos, por ejemplo, qué se entiende por acoso escolar, y las diferencias de género y culturales.

En la mayoría de los casos, las respuestas que hacían referencia a una cantidad reducida de efectos tendieron a demostrar profundidad de conocimiento y obtuvieron una mayor puntuación.

Psicología del deporte

Pregunta 13

Hubo muy pocos alumnos que respondieran a esta pregunta y, aun así, el nivel general de las respuestas no fue muy bueno. Parece que a los alumnos les resultó difícil y fue una pregunta poco popular.

Pregunta 14

Esta fue la pregunta más popular en su opción y atrajo la atención de los alumnos que claramente habían estudiado esta opción. La mayoría de las veces, las respuestas hicieron referencia a la hipótesis de la U invertida, el modelo de catástrofes, la teoría de la activación óptima o la teoría de la autoeficacia.

Los alumnos tuvieron más problemas, incluida la evaluación de la teoría seleccionada, pero algunos lograron abordar correctamente las siguientes cuestiones:

- Consideraciones metodológicas
- La precisión y claridad de los conceptos
- Hallazgos o explicaciones contradictorios
- La productividad de la teoría para generar investigaciones psicológicas
- Las aplicaciones de los hallazgos empíricos

Los alumnos que abordaron una teoría o una cantidad reducida de teorías tendieron a proporcionar mejores respuestas, ya que demostraron conocimiento e indicios concretos de pensamiento crítico.

Pregunta 15

Esta fue una elección bastante popular en esta opción, pero muchas veces la eligieron alumnos que obviamente tenían poco conocimiento específico del tema. La mayoría de los alumnos demostró un conocimiento de sentido común sobre el tema. En muchos casos, los alumnos solamente abordaron teorías o modelos de estrés sin relacionarlos con la respuesta del atleta frente al estrés.

Fueron pocas las ocasiones en que hubo una discusión en profundidad, pero cuando sí sucedió, la discusión incluyó consideraciones culturales, metodológicas y de género.

Recomendaciones y orientación para la enseñanza de futuros alumnos

- En la mayoría de los casos, el principal problema es la incapacidad de interpretar los términos de instrucción. Por ese motivo, es importante que, desde el principio del curso, los alumnos estén familiarizados con los términos de instrucción y estén expuestos a tipos de preguntas parecidos a los que aparecen en los exámenes del IB. Así, estarán bien preparados para los exámenes finales del IB.
- Los alumnos deben comprender la importancia de responder la pregunta que se les plantea. Aportar muchos conocimientos puede resultar impactante, si bien no sirve de nada si no son pertinentes para la pregunta planteada.
- Los profesores deben intentar alentar a los alumnos para que establezcan una conexión entre la teoría, el concepto o el término y los estudios empíricos; además, se deben asegurar de que, al hacerlo, evalúan también el concepto y no solo los estudios empíricos, de acuerdo con los requisitos de la pregunta.
- Los profesores deben instar a los alumnos a centrarse en los aspectos principales de la investigación psicológica, incluidos el nombre de los autores, el tipo de estudio, dónde se llevó a cabo, los aspectos principales del procedimiento y los hallazgos del estudio. Así se garantizará una descripción precisa.

Prueba 3 del Nivel Superior

Límites de calificación del componente

Calificación final:	1	2	3	4	5	6	7
Rango de puntuaciones:	0-1	2-3	4-6	7-9	10-13	14-16	17-30

Comentarios generales

Un número considerable de alumnos pareció haber entendido bien el material de estímulo, y la mayoría de los alumnos se refirió a dicho material en sus respuestas. Sin embargo, en algunas respuestas los alumnos parecieron estar confundidos sobre cómo se utilizaba el diario de campo en el estudio. Algunos alumnos también tuvieron problemas para comprender el papel de la investigadora en las observaciones del participante; de hecho, parece que creían que la investigadora era una paciente o una participante real. Como es habitual, muchos alumnos utilizaron el término “experimento” como término genérico para “estudio”, por lo que tal vez haya que tenerlo en cuenta en el futuro a la hora de enseñar métodos de investigación cualitativos.

Hubo algunas respuestas muy buenas que reflejaron que algunos alumnos estaban muy bien preparados para la prueba 3. Las mejores respuestas demostraron un conocimiento preciso de la metodología de investigación cualitativa; en ellas, los alumnos supieron aplicar eficazmente el material de estímulo utilizando los conceptos y los argumentos adecuados. Esto resultó muy gratificante, ya que demuestra que los alumnos pueden adquirir un conocimiento adecuado de los métodos de investigación cualitativos y aplicarlos eficazmente al material de estímulo.

En general, las respuestas más flojas tendieron a no centrarse en las preguntas planteadas; en lugar de eso, los alumnos comentaron el material de estímulo sin referirse al conocimiento de los métodos de investigación cualitativos. Por ejemplo, los alumnos abordaron la terapia ofrecida, el entorno o las cuestiones éticas que se deben tener en cuenta a la hora de desarrollar una investigación en un hospicio, o bien intentaron interpretar las emociones de los pacientes terminales en vez de abordar el hecho de que el material de estímulo representa un estudio de caso cualitativo.

En general, las peores respuestas utilizaron conceptos y conocimientos pertinentes limitados de los métodos de investigación cualitativos. Igual que en convocatorias anteriores, las respuestas más flojas mostraron cierta tendencia a razonar basándose en la investigación cuantitativa; de hecho, algunos alumnos demostraron que tenían una comprensión muy limitada de los métodos de investigación cualitativos y del estudio de caso, por ejemplo, al referirse a la falta de datos cuantitativos como limitación (en la pregunta 2).

Áreas del programa y del examen que resultaron difíciles a los alumnos

En la pregunta 1, parece que los alumnos se confundieron un poco sobre el hecho de que la pregunta debía centrarse en los posibles efectos de las expectativas de los participantes sobre los datos obtenidos y no en lo que, como pacientes terminales, esperaban obtener de la terapia.

En lo referido a la pregunta 2, parece que en general supuso la mayoría de los problemas para los alumnos, ya que estos no demostraron suficiente conocimiento de un estudio de caso y por qué se usa, o simplemente se limitaron a describir lo que se hacía en el estudio sin abordar realmente la pregunta.

A pesar de que, en general, la pregunta 3 estuvo bien respondida por los alumnos, las peores respuestas a esta pregunta demostraron una tendencia a desarrollar un argumento basado en la suposición de que los investigadores por naturaleza están tan sesgados que quieren expresar sus propias ideas en un estudio. Esto demuestra que, en general, no se comprende que el propósito de una investigación cualitativa es comprender el mundo subjetivo de los participantes; asimismo, refleja que no se comprende que no hay ningún problema con que el investigador tenga un interés personal en el área de investigación, y que la reflexividad es un enfoque para tener en cuenta los sesgos a fin de asegurar que sean las experiencias de los participantes las que se documentan.

Parece que algunos alumnos aún tienden a comentar el estudio en el material de estímulo en vez de abordar explícitamente las preguntas relacionadas con la metodología. Los alumnos más flojos tuvieron problemas para integrar eficazmente el material de estímulo en su respuesta, ya que, o bien confiaron en incluir citas abundantes del material de estímulo en sus respuestas, o bien trataron el material de estímulo como si fuera un texto para analizar en vez de utilizar algunas partes pertinentes para respaldar sus argumentos sobre consideraciones metodológicas y éticas. Algunos alumnos tuvieron problemas con los términos de instrucción y no demostraron comprender cómo abordar la instrucción “explicar” o “discutir”.

Áreas del programa y del examen en las que los alumnos demostraron estar bien preparados

La mayoría de los alumnos utilizó el material de estímulo en cierta medida, y las mejores respuestas integraron el conocimiento pertinente de la metodología de investigación cualitativa con las partes pertinentes del material de estímulo. En lo referido a la pregunta 1, la mayoría de los alumnos demostró algún tipo de conocimiento sobre los efectos de las expectativas del participante en una investigación. En la pregunta 3, fueron muchos los alumnos que demostraron una buena comprensión de la reflexividad y que supieron respaldar su explicación con citas adecuadas procedentes del material de estímulo, por ejemplo, mediante referencias al trabajo de la investigadora y el diseño longitudinal del estudio. Esto refleja que algunos alumnos están bien preparados para detectar detalles pertinentes en el material de estímulo y para utilizarlos de manera adecuada para respaldar su argumento.

Puntos fuertes y débiles de los alumnos al abordar cada una de las preguntas

Pregunta 1

En general, parece que a algunos alumnos les pareció difícil escribir una respuesta con la profundidad suficiente. Si bien resultó obvio que comprendían en cierto modo las expectativas de los participantes, muchos alumnos demostraron una falta de conocimientos precisos.

Las mejores respuestas explicaron qué se entendía por “expectativas de los participantes”; por ejemplo, hubo quien lo definió como “maneras en que los participantes pueden alterar su comportamiento porque saben que forman parte de un estudio”. Estas respuestas también exponían que lo anterior puede afectar a la credibilidad del estudio e incluían ejemplos pertinentes. Por ejemplo, los alumnos con mejores respuestas fueron capaces de explicar que la concesión del consentimiento informado implicaba que los participantes probablemente sabían de qué trataba el estudio, lo cual podía producir sesgos en el estudio de muchas maneras; se hizo referencia al sesgo del participante, el efecto Hawthorne, la reactividad, las características de la demanda, la conformidad o el efecto de la deseabilidad social. Los mejores alumnos se refirieron al carácter público de la observación de los participantes, así como a los procesos sociales en una entrevista de grupo de discusión como factores que podían influir el comportamiento de los participantes.

Los alumnos más flojos tuvieron dificultades para definir las expectativas de los participantes. De hecho, algunos solamente dieron razones irrelevantes de por qué los participantes podrían querer sabotear el estudio. Esto hizo que proporcionaran una descripción superficial del estudio, los participantes y la investigadora sin referirse a la pregunta planteada.

Pregunta 2

Parece que el término de instrucción “discutir” supuso un gran reto en esta pregunta; en general, los alumnos parecieron tener dificultades para responderla. A pesar de que muchos alumnos demostraron algún tipo de conocimiento sobre el estudio de caso o sobre la manera en que se recopilaron los datos en el estudio de caso, en general las respuestas estuvieron poco definidas. La mayoría de los alumnos se refirió a los dos métodos de obtención de datos utilizados. Los mejores alumnos también abordaron el motivo por el cual estos dos métodos podrían ser adecuados en el contexto del estudio de caso en el material de estímulo, por ejemplo, con referencia a los puntos fuertes y las limitaciones de los métodos en relación con el área de investigación del estudio de caso. Algunos alumnos abordaron las características de un estudio de caso; se refirieron al objetivo del estudio como “el caso” y lo desarrollaron más a fondo, demostrando así un conocimiento preciso.

Las mejores respuestas definieron el estudio de caso como un modo de comprender un fenómeno único (el caso), que suele ser de naturaleza delicada (como en este estudio). Las respuestas de este tipo se refirieron a la triangulación, incluyendo referencias a los métodos de obtención de datos (observación pública con participación y entrevistas de grupos de discusión); además, los alumnos discutieron la elección de estos métodos en este estudio de caso en concreto. En muchas ocasiones se discutieron de manera competente los problemas

de la generalización de estudios de caso únicos, por ejemplo, refiriéndose a la posibilidad de la generalización teórica.

Las respuestas más flojas no lograron abordar la pregunta planteada acerca del uso de un estudio de caso, y los alumnos se dedicaron a escribir sobre los métodos utilizados de manera genérica, sin relacionarlos con el material de estímulo. Además, algunos alumnos escribieron sobre las cuestiones éticas o indicaron que la investigadora utilizó una muestra distinta o un método alternativo. Las respuestas más flojas tendieron a repetir lo que se había dicho sobre las expectativas de los participantes en la pregunta 1, o escribieron de manera genérica sobre el estudio de caso y luego lo repitieron.

Pregunta 3

Muchos alumnos demostraron saber al menos un poco sobre la reflexividad como modo de descubrir posibles sesgos en el proceso de investigación. Las mejores respuestas abordaron esta pregunta con explicaciones sobre por qué podía ser importante aplicar la reflexividad en el estudio; además, hicieron referencia a los detalles adecuados en el material de estímulo. Entre estos detalles, se incluía el hecho de que la investigadora trabajaba en el hospicio y conocía a muchos de los pacientes; que se trataba de un estudio longitudinal y que, por ello, podía perder objetividad; y que la investigadora había utilizado varios métodos de obtención de datos y había pedido a un investigador externo que comprobara el análisis de los datos que había hecho. La mayoría de los alumnos se refirió a la reflexividad epistemológica y a la triangulación, así como a la reflexividad personal, haciendo referencias al diario de campo en el que la investigadora había registrado sus propios pensamientos y sentimientos durante la investigación.

Las respuestas más flojas no explicaron qué es la reflexividad y por qué se utiliza en este estudio en concreto. Algunas de estas respuestas también expusieron que el objetivo de la reflexividad es eliminar todos los sesgos, mientras que otros alumnos no entendieron bien el concepto de reflexividad e indicaron maneras de evitarla para que la investigación no estuviera sesgada. Además, las respuestas más flojas tendieron a limitarse a repetir lo que se había dicho sobre las expectativas de los participantes en la pregunta 1, o abordaron de manera genérica la reflexividad.

Recomendaciones y orientación para la enseñanza de futuros alumnos

La prueba 3 se basa en una breve descripción de un estudio de investigación cualitativo (el material de estímulo) acompañado de tres preguntas relacionadas con la metodología utilizada en ese estudio de investigación en particular. Los alumnos deben responder las tres preguntas, prestar atención a los términos de instrucción y utilizar sus conocimientos sobre la investigación cualitativa, así como la información obtenida del material de estímulo para respaldar su análisis. Debe prepararse a los alumnos para que aborden de forma clara cada pregunta y eviten llenar la respuesta de conocimientos generales que no son directamente pertinentes para la pregunta formulada y que, por tanto, no recibirán ningún punto. Lamentablemente, muchos alumnos redactaron párrafos de introducción y de conclusión demasiado largos y que, sin embargo, no aportan demasiado en cuanto a conocimiento y

comprensión. Por ese motivo, es importante enseñar a los alumnos que una pregunta de respuesta corta no exige una introducción y una conclusión.

En general, en la convocatoria de mayo de 2016, los alumnos se refirieron de algún modo al material de estímulo, si bien aún hay algún problema al respecto. Con frecuencia quedó claro que los alumnos comentaron el contenido del material de estímulo utilizando el sentido común, con un conocimiento limitado de los métodos de investigación cualitativos. La mejor manera de prepararse para el examen es utilizar exámenes de convocatorias anteriores para practicar; de este modo, los alumnos tendrán la oportunidad de comprender mejor cómo utilizar el material de estímulo y cómo combinar los puntos pertinentes que proponga este material con un conocimiento y comprensión precisos de los métodos de investigación cualitativos.

Lo ideal es que la enseñanza de la prueba 3 incluya siempre la exposición a varios estudios cualitativos que ofrezcan a los alumnos más oportunidades de comprender la filosofía de la investigación cualitativa. La mejor estrategia es que los alumnos lleven a cabo pequeños proyectos de investigación sobre cada uno de los métodos con el fin de apreciar los razonamientos de un investigador cualitativo, ya que esto les resultaría muy útil para desarrollar las habilidades de pensamiento necesarias para la prueba 3.

Además, se recomienda a los profesores que ofrezcan oportunidades a los alumnos para practicar los términos de instrucción relacionados con la prueba 3. Son demasiados los alumnos que siguen teniendo problemas aquí, por lo que comprender lo que requiere un término de instrucción concreto en la prueba 3 debe formar parte de una enseñanza de calidad.

Finalmente, se recomienda preparar a los alumnos de tal forma que posean (1) un conocimiento general de los métodos de la investigación cualitativa, tal y como se resume en la guía, (2) competencia para aplicar este conocimiento en lo que se refiere al material de estímulo y (3) competencia para emplear en sus respuestas los términos y conceptos adecuados de los métodos de investigación cualitativa. Se recomienda también que se capacite a los alumnos para realizar evaluaciones y discusiones equilibradas, en lugar de presentar opiniones personales y especulaciones que guardan una pertinencia limitada con las preguntas planteadas.