

PSICOLOGÍA TZ2

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

Bandas de calificación de la asignatura

Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 9	10 - 19	20 - 29	30 - 41	42 - 53	54 - 65	66 - 100

Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 10	11 - 23	24 - 33	34 - 45	46 - 56	57 - 67	68 - 100

Variaciones en los exámenes según zonas horarias

Para proteger la integridad de los exámenes, cada vez se está haciendo más uso de versiones distintas de los exámenes según sus zonas horarias. Mediante el uso de variantes de la misma prueba de examen, los alumnos de una parte del mundo no siempre harán el mismo examen que los alumnos de otras partes del mundo. Se aplica un proceso riguroso para garantizar que los exámenes son comparables en términos de dificultad y cobertura del programa de estudios, y se toman medidas para garantizar que se aplican los mismos estándares de calificación a las diferentes versiones de las pruebas de examen de los alumnos. Para la convocatoria de exámenes de mayo de 2012, el IB ha realizado versiones distintas de los exámenes de la prueba 1 de Psicología según sus zonas horarias. Las bandas de calificación para cada uno de los exámenes de distintas zonas horarias se preparan por separado, y se toman las decisiones necesarias en base a los criterios de nivel de rendimiento, para tener en cuenta las diferencias en los exámenes.

Evaluación interna del Nivel Superior

Bandas de calificación del componente

Nivel Superior

Calificación final:	1	2	3	4	5	6	7
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 adecuación del trabajo entregado fue de un buen nivel, a pesar de que este año también se entregaran estudios no experimentales. También se han visto trabajos que no respetaban las pautas éticas, p. ej., las réplicas de los estudios de Asch u otros sobre la conformidad. En la guía de Psicología se indica claramente que dichos experimentos son éticamente inaceptables. La mayoría de los alumnos estaban al tanto de los problemas éticos y la mayor parte de ellos incluyeron una copia del consentimiento informado en los apéndices, aunque unos pocos alumnos no obtuvieron el consentimiento de los padres para los participantes menores de 16 años.

La mayoría de los trabajos se basaban en la psicología cognitiva y esto parece proporcionar buenos resultados en este nivel de educación. Los experimentos más utilizados fueron, como de costumbre, los niveles de procesamiento; la curva de serie-posición; Stroop; Loftus y Palmer (1974), y los experimentos relativos a la teoría de los esquemas mentales. También se vieron varios experimentos basados en el efecto Mozart. Como es habitual, no estuvieron bien hechos, y esto se debe probablemente a que los antecedentes teóricos y empíricos para el efecto Mozart son pobres y a que los hallazgos originales no se han replicado nunca. Muchos alumnos hicieron experimentos con varias condiciones, incluso hasta cinco condiciones. Eso no es un experimento sencillo, como se indica en la guía. No hay razón alguna para tener más de dos condiciones y el empleo de más condiciones no aumenta necesariamente la calidad del trabajo. De hecho, los cálculos estadísticos con frecuencia resultaron erróneos con muchas de las condiciones y la presentación de los resultados fue menos clara.

Por lo general, los peores informes compartieron las siguientes características:

- La explicación de la investigación de contexto de la introducción era pobre e imprecisa. La hipótesis no estaba justificada y operativizada de forma clara.
- Los resultados no estuvieron siempre claramente relacionados con el objetivo del estudio y los tests inferenciales no estaban justificados o brillaban por su ausencia. Por lo que se refiere a la estadística descriptiva, el gráfico presentaba juntas la desviación típica y la media.
- Las discusiones eran superficiales, no discutían los resultados 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.
- Las referencias eran pobres.

Desempeño de los alumnos con relación a cada criterio

La mayor parte de los alumnos parecían estar familiarizados con los requisitos de cada criterio de evaluación, aunque se dieron diferencias en los niveles de logro. Las secciones de introducción estaban, en algunos casos, muy bien escritas, con un enfoque claro, aunque parece que a algunos alumnos les resultaba difícil prestar un enfoque exclusivo a la relación

entre un estudio de investigación concreto y la propia hipótesis de investigación del alumno. La introducción es importante ya que presenta los fundamentos del experimento del alumno y utiliza la investigación de contexto para justificar la hipótesis de investigación del alumno. Por tanto, la investigación de contexto debe explicarse y analizarse con suficiente profundidad (por ejemplo, objetivos, procedimiento, resultados) para permitir la formulación de una clara hipótesis de investigación en la introducción y para estimular la discusión de los resultados del alumno a la luz de la investigación de contexto en la sección de discusión. El experimento debe ser simple y, por tanto, es correcto hacer una réplica parcial de un estudio de investigación, reduciendo, por ejemplo, el número de variables. Las secciones de introducción y discusión son con frecuencia las más difíciles de escribir, ya que requieren una buena comprensión de la vinculación de los estudios de investigación con la formulación de nuevas hipótesis. En ocasiones, el nivel de profundidad del análisis de la investigación de contexto fue superficial, lo que influyó en la discusión, ya que las comparaciones de los resultados del alumno y los resultados de los estudios de contexto no pudieron llevarse a cabo con la suficiente profundidad. Las referencias no tuvieron siempre un formato estándar y sigue habiendo problemas con los recursos encontrados en Internet.

Criterio A

Por lo general, es necesario prestar más atención a la investigación de contexto en los informes. Las teorías y estudios a menudo no se explican con el detalle suficiente para justificar las hipótesis. El objetivo del estudio no se formuló claramente en todos las pruebas. Unas cuantas introducciones incluyeron explicaciones redundantes sobre la investigación que no eran especialmente pertinentes para el estudio del alumno. El análisis de la investigación de contexto se hizo muy bien en algunos casos, aunque resultara superficial en ocasiones y se basara, por ejemplo, en el resumen de un estudio encontrado en un sitio web o en una guía de asignatura. Esto ocasionó algunos problemas a la hora de justificar y formular hipótesis operativizadas de forma clara. Parece que a algunos alumnos les resultó difícil formular una hipótesis del experimento clara y justificada.

Es importante que la introducción sea clara y que esté bien centrada en la investigación de contexto pertinente, de forma que conduzca de manera lógica hacia la hipótesis de investigación del alumno, y que la investigación de contexto se explique y analice con suficiente profundidad (por ejemplo, objetivos, procedimiento, resultados) para permitir la formulación de una clara hipótesis de investigación en la introducción y para estimular la discusión de los resultados propios a la luz de la investigación de contexto en la sección de discusión. Debe tenerse en cuenta que no se exige hacer una réplica exacta de un experimento. Una réplica parcial servirá, pero el experimento del alumno debe estar estrechamente relacionado con un experimento real y no sencillamente inspirado por cierto tipo de investigación. Resultó evidente que algunos alumnos habían elegido este enfoque y que escribieron, por tanto, introducciones muy pobres.

Criterio B

La mayoría de los alumnos parecían conocer lo que se entiende por el diseño experimental, pero la elección del diseño no siempre se justificó debidamente, como,

por ejemplo, haciendo referencia a los puntos fuertes y débiles de los diseños respectivos (de medidas repetidas e independientes). Algunos alumnos tuvieron dificultades para la operacionalización de las variables dependiente e independiente. Las normas éticas se abordaron en su mayor parte en la sección de diseño, aunque a veces se hiciera en los procedimientos, lo cual es correcto. Lo que es fundamental es que los procedimientos éticos se aborden de una forma adecuada en algún lugar del informe. La mayor parte de los alumnos incluyeron un consentimiento informado de los participantes. Cierta número de colegios habían permitido la participación de menores de 16 años sin el consentimiento de sus padres, lo que constituye una violación de las pautas éticas para la evaluación interna. Tan solo hubo unos pocos casos en los que no se incluyó el consentimiento informado, y esto siempre se dio en informes no muy buenos en los que también faltaron otras cosas.

Criterio C

A menudo hubo problemas en la identificación de las características pertinentes de los participantes. Es posible que sea difícil definir las características pertinentes respecto de un estudio concreto, pero los alumnos deben centrarse en características como la edad, el sexo, el número de participantes y la nacionalidad. No siempre se identificó la población objetivo. Todas las muestras se extrajeron del alumnado del colegio del propio alumno. La técnica de muestreo se identificó claramente en la mayoría de los casos como muestra de conveniencia o muestra autoseleccionada y se explicó o justificó a menudo. Ambos enfoques son correctos. Algunos alumnos dijeron haber utilizado una muestra aleatoria, lo que claramente no fue el caso. Parece haber cierta confusión sobre el significado de la palabra "aleatorio". Algunos alumnos no especificaron cómo habían asignado los participantes a las condiciones experimentales, pero no se les puede penalizar por ello porque no se menciona en los criterios de evaluación.

Criterio D

Con bastante frecuencia ha habido problemas con el grado de detalles de la descripción del procedimiento, lo que hace difícil replicar un procedimiento. No siempre se hacía referencia a los materiales en los apéndices (bien sea en el procedimiento o en los materiales), y en esos casos la replicación sencillamente no era posible. Se recomienda incluir todos los materiales y dar detalles de cómo se utilizaron en el procedimiento.

Criterio E

La mayoría de los alumnos describieron los resultados de manera narrativa en la sección de resultados. No todos incluyeron la desviación típica como estadística descriptiva, ni aunque sus datos lo permitieran. Algunos utilizaron el rango, pero no explicaron su uso. Hay cierta tendencia a incluir toda una serie de medidas de tendencia central en la sección de resultados, pero esto es redundante. No afecta a las puntuaciones, pero es un indicio claro de que los alumnos no han tenido en cuenta qué medida sería la más apropiada para sus datos.

En algunos informes no había gráfico alguno, sino tan solo una tabla de resultados. A menudo los gráficos no estaban muy bien rotulados. No todos los informes incluyeron tablas. Unos pocos alumnos incluyeron resultados individuales en la sección de resultados, si bien la mayor parte de los alumnos colocaron correctamente los datos sin procesar en los apéndices. Bastantes de los alumnos representaron gráficamente su desviación típica junto a la media, lo que no tiene realmente sentido y debe evitarse.

Criterio F

La mayor parte de los alumnos eligieron y justificaron el test de estadística inferencial correctamente, aunque unos pocos no lo hicieron así. La mayoría de los alumnos utilizaron los tests no paramétricos, aunque un número cada vez mayor de alumnos utilizará la prueba t, que también es apropiada dado que esos tests son bastante sólidos. No todos los alumnos pudieron justificar la prueba estadística que habían elegido y algunos de ellos no hicieron un enunciado de significación estadística.

Criterio G

Varios de los alumnos que lograron las mejores calificaciones demostraron su competencia en la discusión de sus propios resultados a la luz de la investigación de contexto. Esos alumnos, además, identificaron las limitaciones metodológicas y sugirieron las modificaciones pertinentes. Sin embargo, muchos no fueron capaces de integrar adecuadamente la investigación de contexto de la introducción en la discusión de sus propios resultados. Muchos informes presentaron una discusión muy corta y superficial de sus resultados a la luz de la investigación previa y se limitaron a una larga descripción de las limitaciones del propio diseño, limitaciones que no siempre resultaban pertinentes para el estudio sino más bien una lista de carácter más general. Algunos alumnos incluyeron los puntos fuertes de su diseño, a pesar de que ya no es necesario. Las sugerencias para las modificaciones estaban a menudo relacionadas con los factores generales, en lugar de con el estudio del alumno. Por ejemplo, hubo muchos alumnos que mencionaron la muestra puntual como un problema y sugirieron una muestra aleatoria y más participantes para futuras investigaciones.

Criterio H

El problema más común gira en torno a las referencias. A menudo los alumnos no incluyeron todas las referencias que mencionaban en la introducción. Sigue habiendo problemas con las referencias procedentes de Internet, ya que los alumnos suelen pensar que basta con la URL y en ocasiones no se puede encontrar el estudio que se presenta como contexto en la sección de referencias. Las referencias no seguían a menudo un método estándar de referencias, como el de la APA, o no estaban completas. Por lo general, las referencias secundarias eran pobres.

Criterio I

Algunos alumnos perdieron puntos aquí porque les faltaban partes del resumen o de los apéndices

Recomendaciones para la enseñanza a futuros alumnos

- Es muy importante ayudar a los alumnos a encontrar la investigación de contexto y el marco teórico adecuados, ya que dichos recursos les permitirán analizar la investigación de contexto con cierta profundidad. Se recomienda encontrar experimentos relativamente sencillos de replicar. Esto ayudaría a los alumnos a hacer por sí mismos experimentos relativamente sencillos basándose en experimentos científicos reales. Para los alumnos es mucho más fácil replicar un experimento si tienen acceso a lecturas sobre los experimentos o resúmenes de ellos, de forma que puedan leer sobre las hipótesis y otros detalles importantes.
- Se debe centrar más la atención en la relación entre los objetivos del estudio del alumno y la investigación de contexto, de forma que ambas se puedan integrar en la introducción y la discusión de los resultados. La investigación de contexto se debe analizar con la suficiente profundidad en la introducción como para que queden justificados claramente los objetivos de la investigación del alumno, y la hipótesis del experimento debe estar estrechamente relacionada con la investigación de contexto.
- La sección de diseño debe incluir una descripción clara de las condiciones experimentales y los profesores deben intentar limitar las condiciones experimentales a dos solamente (bien dos condiciones de tratamiento o una condición de tratamiento y una de control). Esto haría más fácil comparar los resultados de la manipulación de la variable independiente en la variable dependiente en las dos condiciones. Esto sigue también la recomendación del IB de realizar estudios experimentales sencillos con los alumnos de Psicología.
- El muestreo debe hacerse de acuerdo con las normas del IB, es decir, la identificación de la población objetivo, incluidas las características pertinentes y la descripción del método de muestreo, así como la explicación (o justificación) del uso del método que se ha elegido. La mayoría de los alumnos utilizaron una muestra de conveniencia, pero aún así deben explicar el método de muestreo o justificar su uso. No es necesario que el número de participantes en el experimento sea mayor de 20 (diseño independiente) o 10 (diseño de medidas repetidas).
- Los gráficos y tablas deben tener un título adecuado en la sección de estadística descriptiva. Debe resaltarse que la representación gráfica de los resultados es obligatoria y que debe incluirse una tabla. Esta sección incluye datos *resumidos*, no

datos sin procesar o resultados individuales. Se recomienda a los alumnos que no incluyan varias medidas de tendencia central, sino solamente la que sea expresamente pertinente para sus datos. Del mismo modo, debe calcularse una medida apropiada de dispersión y deben incluirse los resultados en esta sección.

- En la sección de estadística inferencial, los alumnos deben tener cuidado a la hora de elegir un test estadístico adecuado y deben justificar por qué eligen ese test. Esto puede relacionarse con el nivel de medición de datos y el diseño del experimento.
- En la sección de discusión debe hacerse referencia a la explicación de los estudios empíricos y al marco teórico de la introducción. No deben incluirse aquí los nuevos estudios o teorías. Los alumnos deben tener la preparación necesaria para elaborar una sección de discusión (tal vez mediante la lectura de un par de artículos de investigación para familiarizarse con la idea y el estilo) y deben consultar la lista de verificación para asegurarse de que se cumplen todos los requisitos del IB. La comprensión de las limitaciones del procedimiento y las sugerencias de modificación deben estar unidas. No es suficiente decir que un estudio concreto debe utilizar un muestreo aleatorio e incluir a más participantes para que sea mejor. Las limitaciones deben ser claramente pertinentes para el experimento del alumno.
- Se recomienda en general que los alumnos se familiaricen con los estándares científicos, que incluyen referencias a estudios anteriores, y la integración de estos en la propia investigación. Del mismo modo, debe fomentarse el uso de lecturas de apoyo 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 especialista como bibliografía de apoyo. Además, los alumnos pueden obtener más puntos en el criterio H si aprenden a utilizar un método estándar de referencias.

Evaluación interna del Nivel Medio

Bandas de calificación del componente

Nivel Medio

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

Ámbito y adecuación del trabajo entregado

En conjunto, el ámbito y adecuación del trabajo entregado fue de un nivel satisfactorio y este año se enviaron menos estudios no experimentales. Los temas más populares fueron de la psicología cognitiva, y esos temas suelen ser apropiados y manejables para alumnos de este nivel de educación. El efecto Stroop y los experimentos de Loftus y Palmer fueron con frecuencia los más replicados. Algunos alumnos prefirieron ocuparse del efecto de primacía y recencia y tuvieron dificultades para explicar sus resultados de manera adecuada. A los moderadores les ha encantado observar que la mayoría de los alumnos del Nivel Medio están llevando a cabo ahora experimentos que son lo bastante sencillos como para alcanzar los objetivos de la evaluación interna y no sobrepasar el límite de palabras. Sin embargo, se han dado ejemplos de alumnos que comparaban tres o más variables, cuando solo necesitaban dos. Esto ha ocurrido a menudo cuando los alumnos replicaban el experimento de Loftus y Palmer sobre las preguntas que inducen ciertas respuestas.

Los moderadores informaron de que la mayoría de los alumnos eran conscientes de las consideraciones éticas básicas. La mayoría de los trabajos contenían pruebas apropiadas, lo que indica que los alumnos conocían y respetaron los principios éticos durante la planificación, durante la realización de la investigación y al elaborar el informe sobre sus hallazgos. Sin embargo, algunos alumnos siguen sin conseguir el consentimiento de los padres para los participantes que son menores de 16 años.

Por lo general, los moderadores informaron de una mejora en el estándar general de los informes. En general, la introducción se hizo bastante bien. Parece que aumenta el número de alumnos que consiguen ahora todos los puntos en este criterio. Además, los aspectos relacionados con los procedimientos estuvieron por lo general bien hechos y se informó bien de ellos; sin embargo, se sigue obteniendo malos resultados en el criterio B, que pide elegir, justificar y explicar el diseño, y el criterio E, que pide analizar los datos. Normalmente, la discusión de los resultados a la luz del estudio original fue muy superficial.

Por lo general, los peores informes compartieron las siguientes características:

- Eligieron un tema inapropiado (por ejemplo, los objetivos generales que hacían referencia a las diferencias de sexo o los objetivos generales diseñados únicamente por el alumno sin proporcionar un estudio original).
- Se identificaron el diseño de la investigación y el método de muestreo, pero no se justificaron.
- No se proporcionó una justificación pertinente para las estadísticas descriptivas.
- Las discusiones carecían de información pertinente y se presentaban de una manera superficial.
- No se indicaba el número de palabras y no se respetaba el número máximo de palabras establecido.

- No se proporcionaba una página de referencias o se presentaban las referencias sin información suficiente.

Se observaron los siguientes problemas en algunos de los temas elegidos que son populares:

- Cuando replicaban el estudio de Loftus y Palmer (1974) sobre la memoria reconstructiva, los alumnos comparaban con frecuencia tres o más variables, cuando solo necesitaban dos. Debe recordarse a los alumnos que utilicen dos condiciones, por ejemplo, “estrellaron” y “tocaron”. Además, los alumnos que replicaron el estudio de Loftus del “accidente de tráfico” siguen sin incluir la referencia al “video de los accidentes de tráfico” en el formulario o escrito de consentimiento, y los profesores deben recordarles que tienen que hacerlo, en lugar de proporcionar a toda la clase un formulario de consentimiento genérico. El objetivo de los formularios de consentimiento es informar sobre el estudio para permitir la retirada de los participantes.
- Cuando se replicaba el efecto Stroop, un número significativo no explicó bien la variable independiente de palabras coloreadas que eran congruentes y no congruentes; a menudo los alumnos utilizaron explicaciones o términos distintos en el mismo informe. Muchos centros enviaron fotocopias de los informes en lugar de los originales, lo que fue con frecuencia un problema para distinguir bien los materiales incluidos en el apéndice.
- El estudio de Bartlett, *La guerra de los fantasmas*, fue replicado con frecuencia, aunque no haya una manipulación clara de una variable independiente.
- Algunos alumnos llevaron a cabo una réplica del experimento de Asch sobre la conformidad, aunque este estudio presente consideraciones éticas serias.

Desempeño de los alumnos con relación a cada criterio

En su mayor parte, se hizo evidente una amplia distinción entre los alumnos en cada criterio.

Criterio A

En general, la introducción se hizo bastante bien. Muchos alumnos proporcionaron introducciones claras y no les resultó difícil conseguir todos los puntos. Sin embargo, muchos alumnos siguen con frecuencia incluyendo información que no era necesaria en la introducción: teorías y estudios adicionales e hipótesis de investigación e hipótesis nulas que no eran necesarias y ocuparon un número valioso de palabras. Otros problemas adicionales que indicaron los moderadores fueron que el objetivo general no siempre se indicaba claramente o no coincidía con la descripción del estudio original. En algunos casos, la formulación del objetivo general eran imprecisa y solo estaba tangencialmente relacionada con el estudio que se estaba replicando, con unas explicaciones mínimas sobre las modificaciones que se habían implementado.

Criterio B

La sección de diseño parece ser un punto débil de muchos informes. Además, algunos elementos contenidos en este criterio están relacionados con otras secciones del informe y criterios (criterio E y criterio F), y si no se entienden bien, esto tiene consecuencias más profundas cuando se evalúa la calidad del informe. Muchos moderadores informaron de que los alumnos omitieron con frecuencia la mitad de los requisitos o no entendieron cómo explicarlos. En muchos casos, aunque se indicó el diseño de la investigación, no se justificó su uso. En unos pocos casos, la explicación de las medidas repetidas o de las muestras independientes parecía indicar que los alumnos no comprendían los conceptos, a pesar de que utilizaban la terminología apropiada. Sin embargo, los alumnos sí parecían ser capaces de indicar coherentemente las variables dependientes e independientes, mucho mejor que en años anteriores. En ocasiones, aunque se identificaran las variables, no se expresaba su valor con exactitud (por ejemplo, cómo se medía la variable: segundos, minutos u horas para el tiempo y kilómetros o millas para la velocidad)

Normalmente, las normas éticas se han seguido de forma clara y se han proporcionado pruebas de ello dentro del informe. La mayoría de los alumnos adjuntaron una copia en blanco del formulario de consentimiento y la carta de *debriefing* en los apéndices. A algunos se les olvidó indicar las consideraciones éticas en esta sección. Sin embargo, si hacían referencia a las consideraciones éticas en la sección de participantes o en la sección del procedimiento, los moderadores tenían instrucciones claras de conceder los puntos bajo este criterio.

Criterio C

Muchos alumnos eligieron a alumnos de su clase o colegio como participantes. La elección es apropiada y completamente adecuada para sus estudios. Algunos alumnos obtuvieron muestras muy grandes en sus estudios; el tamaño adecuado que se recomienda es de 15 a 20 participantes y no hay necesidad alguna de grupos más grandes de participantes.

A menudo hubo problemas en la identificación de las características pertinentes de los participantes. Además, aunque normalmente se proporcionaba un método de selección, no llegaba a explicarse por completo. Otro problema que se han encontrado muchos moderadores es que los alumnos siguen teniendo dificultades para entender la diferencia entre los conceptos de “selección aleatoria” y “asignación aleatoria”. El término “aleatorio” sigue sin ser entendido correctamente por muchos alumnos, lo que lleva a confusos relatos sobre la selección de los participantes que han utilizado en sus estudios; algunos alumnos tienden a describir con frecuencia las muestras puntuales o de conveniencia como muestras aleatorias o a utilizar el muestreo aleatorio como un equivalente de “cualquiera disponible”.

Criterio D

La mayoría de los alumnos ofrecieron una información clara y pertinente en la sección de procedimiento. Sin embargo, las instrucciones estandarizadas no eran siempre claras o no se incluían en los apéndices.

Muchos alumnos siguen escribiendo procedimientos muy breves y genéricos, que pueden referirse a varios experimentos (como repartir formularios de consentimiento, seguir instrucciones estandarizadas, comparar los resultados y otros parecidos), y luego se apoyan en instrucciones estandarizadas detalladas en los apéndices. Debe proporcionarse más información específica en la sección de procedimiento para indicar las características principales del procedimiento del estudio.

Criterio E

Esta parte del informe tuvo redacciones muy pobres y parece que hay mucha confusión sobre lo que es necesario. En la tabla debe proporcionarse una medida de dispersión y una medida de tendencia central para la condición experimental y la condición de control. Estas medidas deben ser apropiadas para el nivel de medición de datos y para la distribución de los datos que se haya obtenido. Desgraciadamente, muchos alumnos no emparejaron estas medidas de forma apropiada: por ejemplo, los alumnos utilizaron la media, pero luego usaron el rango como medida de dispersión en lugar de la desviación típica, o utilizaron la mediana y luego incluyeron la desviación típica en lugar del rango intercuartil. Además, muchos alumnos siguen incluyendo las tres medidas de tendencia central sin explicación alguna. En general, la justificación del uso de una estadística descriptiva concreta continúa siendo difícil para los alumnos. Muchos alumnos presentaron datos sin ningún texto verbal y, con frecuencia, incluyeron datos sin procesar en la sección.

Los gráficos también deben presentar dos medidas de tendencia central (una para el grupo experimental y otra para el grupo de control) y deben incluirse etiquetas claras y un título. Desgraciadamente, los gráficos y tablas a menudo comparaban datos mal elegidos que no guardaban relación con las condiciones de la variable independiente y el objetivo general.

Criterio F

Las secciones de discusión siguen siendo las peores, y se concedieron muy pocas calificaciones altas. Normalmente, las discusiones de los resultados a la luz del estudio original fueron muy superficiales. Los alumnos no compararon las diferencias en los cálculos y fue escasa la discusión sobre por qué los hallazgos del estudio replicado eran similares o diferentes al estudio original. Una mayor evaluación de las limitaciones del diseño y del procedimiento de investigación produciría unas calificaciones mejores en esta área. Desgraciadamente, algunos alumnos no indicaron las modificaciones ni proporcionaron una conclusión a su sección de discusión, por lo que solo obtuvieron un número mínimo de puntos.

Varios de los alumnos hablaron una vez más de limitaciones generales y superficiales, como el tamaño de la muestra y clases ruidosas. Al mismo tiempo, parece que no eran conscientes de las graves variables confusas que se produjeron.

Otro problema habitual fue que aunque se proporcionó parte de la información pertinente en la sección de discusión, no se desarrolló con todo detalle proporcionando una relación clara entre la metodología utilizada en el estudio y los resultados obtenidos. La evaluación fue con frecuencia implícita o inapropiada.

Criterio G

Las tres mayores limitaciones relacionadas con el criterio G fueron que los alumnos no respetaron el límite de palabras; que no proporcionaron los cálculos estadísticos en los apéndices, o que no incluyeron una página de referencias.

Varias de las muestras excedían claramente el límite de palabras y algunas incluso lo anunciaban en la portada; aún así, el profesor les concedió la puntuación máxima.

Los apéndices se identificaron de un modo apropiado (en la mayoría de los casos) y se hacía referencia a ellos en el informe. Muchos alumnos experimentaron problemas con las referencias, que a menudo no se llevaron a cabo bien o no se hicieron de forma muy sistemática. Las referencias a las fuentes de Internet, en especial, no fueron muy claras. Algunos alumnos tuvieron problemas con la redacción de unos resúmenes claros y concisos. Algunos moderadores informaron de problemas en la presentación del informe: en algunos casos se vieron problemas de ortografía y errores gramaticales sencillos o una presentación e impresión de poca calidad. Debe aconsejarse a los alumnos que envíen informes originales (no copias en blanco y negro) cuando tengan gráficos con colores en la sección de resultados.

Recomendaciones para la enseñanza a futuros alumnos

- Se necesita ayuda para seleccionar los temas apropiados y la investigación a replicar.
- En muchos centros sigue siendo necesario que los profesores expliquen de forma más clara lo que se espera que se proporcione en la evaluación interna a los alumnos. Los profesores deben revisar los criterios de revisión con los alumnos para asegurarse de que están al tanto de lo que se espera de ellos. Hay diferencias bastante notables entre las muestras de centros donde es evidente que los profesores se tomaron el tiempo de analizar los criterios de evaluación interna y aquellos que entraron en menos detalles en la discusión sobre las expectativas.
- Se recomienda a los profesores que den a los alumnos la oportunidad de calificar una prueba de muestra cuando enseñen a alumnos en el futuro. Las discusiones en grupo también mejoran la comprensión de los alumnos de qué elementos deben incluirse en cada sección y la mejor forma de escribir el informe.
- La creación de actividades breves y simples que animen a los alumnos a sopesar los puntos fuertes y débiles de los diferentes diseños y técnicas de muestreo puede ayudar a la comprensión de por qué se utiliza un diseño en concreto y por qué pueden utilizar una muestra de conveniencia en el entorno de un colegio. La comprensión de esos conceptos permitirá a los alumnos justificar mejor la elección

de un diseño de investigación y explicar bien qué tal les fue con la selección de una muestra entre los alumnos disponibles.

- Una exhaustiva presentación y discusión de los tipos de datos (nominales, ordinales, de intervalo y de razón) hubieran sido útiles para enseñar a los alumnos cuándo es apropiado utilizar el modo, la mediana y la media. Además, la creación de histogramas de frecuencia con datos sin elaborar permitirá a los alumnos ver si los datos son asimétricos o reflejan una distribución normal. Este tipo de actividad puede ayudar a los alumnos a entender cómo deberían justificar su uso de una medida de tendencia central en concreto. Además, sería beneficioso enseñarles qué medida de dispersión coincide mejor con una medida de tendencia central. El empleo de más discusión y actividades utilizando datos ficticios podría ayudar a los alumnos a comprender los diferentes tipos de datos y a elegir la estadística descriptiva adecuada cuando enuncien sus resultados.
- Concretamente, deben explicarse a los alumnos con mayor claridad las expectativas para la sección de discusión. Son demasiados los alumnos que no logran abordar los elementos más básicos (resumir sus hallazgos, hacer comparaciones con el estudio anterior, revisar la metodología –elección del diseño y otras opciones–). Hay demasiados comentarios en la discusión que son superficiales.
- Tal vez pueda resultar útil a los profesores que creen ejercicios que ayuden a los alumnos a evaluar datos del estudio y datos originales del estudio replicado del alumno, y luego hacer una lluvia de ideas para analizar por qué fueron similares o diferentes los datos. Los alumnos desarrollarán formas de mejorar sus discusiones si indican los puntos fuertes y las limitaciones de sus propios diseños experimentales y procedimientos. Además, es imprescindible que los alumnos sugieran modificaciones de las áreas que son problemáticas en su experimento y que proporcionen una conclusión básica de sus hallazgos en el experimento.
- Las referencias deben enseñarse de forma explícita, y deben ser bastante coherentes en cada colegio.
- Debe insistirse a los alumnos para que utilicen los correctores de ortografía y de gramática antes de entregar su trabajo final.
- Deben respetarse los límites de palabras y el profesor debe comprobarlos.
- Los profesores deben alentar unos hábitos de lectura más amplios. Los mejores proyectos fueron aquellos en los que era evidente que los alumnos habían leído más cosas que el estudio que estaban intentado replicar.
- Debe insistirse a los alumnos para que se aseguren de que el estudio que están llevando a cabo y del que están dando cuenta es experimental; es decir, que contiene una manipulación clara de una variable independiente. Un cuestionario puede formar parte del estudio si se utiliza para medir la variable dependiente, pero debe prestarse una atención especial para garantizar que el alumno manipula claramente la variable independiente. En los trabajos en los que se utilizan los

cuestionarios como medida de la variable dependiente, los profesores deben animar a los alumnos a considerar cuál es la formulación más apropiada de las preguntas en los cuestionarios.

- Debe recomendarse a los alumnos que no elijan a participantes menores de 16 años para sus estudios.

Prueba 1 del Nivel Superior y del Nivel Medio

Bandas de calificación del componente

Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 5	6 - 10	11 - 14	15 - 19	20 - 24	25 - 29	30 - 46

Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 5	6 - 10	11 - 14	15 - 19	20 - 24	25 - 29	30 - 46

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

La comprensión de las exigencias de las preguntas sigue siendo un problema para muchos alumnos. Muchos de ellos aplicaron una estrategia única para todas las preguntas, sin que importara el término de instrucción. Esto fue así en la redacción de los objetivos generales del nivel de análisis, en la redacción de la historia del nivel de análisis y, en el caso de las preguntas de respuesta corta de la sección A, en la evaluación de la investigación. Los alumnos deben saber que estas estrategias llevan a respuestas mal enfocadas. Además, con frecuencia se escriben varios ejemplos cuando solo se pide uno. Es importante que los alumnos sepan que solo se evalúa la primera respuesta.

Para las preguntas de respuesta corta, los alumnos con frecuencia discutían un concepto y una investigación (un estudio o una teoría), sin relacionarlos entre sí. Por ejemplo, en la pregunta 2 se describieron una consideración ética y un estudio, pero no se indicó ningún vínculo, y en la pregunta 3 se describieron un principio y un ejemplo de un estudio de investigación, pero no quedaba clara la vinculación del principio a la investigación.

Muchos alumnos proporcionaron una evaluación de la investigación superficialmente pertinente o mal desarrollada en las respuestas de desarrollo. Los alumnos deben clarificar su evaluación vinculándola directamente con el estudio. Por ejemplo, algunos estudios recibieron críticas por ser “ecológicamente” válidos (o no), sin hacer ningún intento de explicar lo que esto significa en lo que se refiere al estudio.

Pocos fueron los ensayos que contenían una introducción clara a la pregunta y los conceptos claves solo se definían de vez en cuando. En muchos casos, la conclusión era superficial y los alumnos no hacían relación a la pregunta o al término de instrucción. En otros casos, los ensayos se redactaron en un solo párrafo, sin una transición clara entre las ideas y el desarrollo de un argumento.

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

El conocimiento general de los tres niveles de análisis fue satisfactorio, y la mayoría de los alumnos fue capaz de indicar las explicaciones teóricas y los conceptos adecuados, y de proporcionar ejemplos de investigación pertinentes.

Puntos fuertes y débiles de los alumnos al abordar las distintas preguntas

Sección A

Nivel de análisis biológico

Se han dado algunas respuestas excelentes esta pregunta. Los alumnos han demostrado una buena diversidad de investigaciones, incluyendo los estudios de Maguire, Thompson, Corkin y LeVay. Dicho eso, han sido muchos los alumnos que han intentado describir el estudio de caso de Phineas Gage. Ese estudio es muy antiguo y los alumnos a menudo proporcionaron descripciones inexactas de la investigación y de la localización real de las funciones. Se recomienda encarecidamente a los profesores que ofrezcan a los alumnos investigaciones más modernas que se puedan discutir en el marco de la comprensión psicológica actual.

Las descripciones de los estudios con frecuencia carecían de unos objetivos generales, un procedimiento o unos resultados identificados claramente. A menudo las respuestas se centraban en las consideraciones éticas o en cuestiones de generalización, que eran irrelevantes para la pregunta y con frecuencia incorrectas en su interpretación.

Nivel de análisis cognitivo

Un número sorprendente de alumnos mostró un conocimiento limitado de las consideraciones éticas. Muchas de las respuestas eran muy superficiales y no explicaban con precisión el estudio. Un número significativo de alumnos no describió con detalle las consideraciones éticas. Algunos alumnos utilizaron “asuntos futurísticos” que eran especulativos y no muy pertinentes. Además, afirmar únicamente que la “investigación animal” es una consideración ética no fue una respuesta válida.

Fue muy habitual que los alumnos resumieran varias consideraciones éticas en lugar de centrarse en una, lo que dio como resultado calificaciones bajas. Varios de los alumnos discutieron estudios que no eran cognitivos, incluido el estudio de Milgram

sobre la obediencia. Es importante que los alumnos relacionen de forma clara el estudio con el nivel cognitivo de análisis, identificando el proceso cognitivo que se está estudiando.

Nivel de análisis sociocultural

Muchos de los alumnos describieron tres o cuatro principios del nivel de análisis, lo que hizo que solo se evaluara la primera descripción. Con frecuencia solo se relacionaba de manera superficial el principio y el estudio. Se han dado respuestas excelentes referidas al principio de “nuestro comportamiento se ve afectado con frecuencia por otras personas, incluso cuando creemos que estamos actuando de manera independiente”.

Muchos alumnos incluyeron una teoría y un estudio, aunque solo fuera necesario un ejemplo de investigación. De nuevo, solo se ha evaluado el primero.

Sección B

Nivel de análisis biológico

Se han dado varias respuestas muy buenas a esta pregunta. Muchos alumnos discutieron la interacción entre la cognición y la fisiología por lo que se refiere a la emoción, la amnesia y el estrés. Se ha visto un buen uso de la investigación para describir esta interacción. Ha habido respuestas excelentes que se han centrado en la meditación, pero también ha habido algunas respuestas en las que no era evidente la discusión de los factores cognitivos.

Han sido varios los alumnos que no entendieron las exigencias de la pregunta y compararon o contrastaron un argumento biológico y cognitivo para el origen del comportamiento anormal o para los orígenes de la atracción.

Los alumnos con frecuencia no entendieron la diferencia entre la investigación experimental y la investigación correlacional. Además, muchos alumnos sostuvieron que la investigación biológica no puede generalizarse si no se estudian las diferentes etnias o razas. Este es un argumento problemático. Los alumnos deben ser más conscientes de los puntos fuertes y las limitaciones de la investigación biológica.

Por lo que respecta al aspecto del pensamiento crítico de la pregunta, los alumnos a menudo solo evaluaron la metodología de los estudios presentados, sin discutir la interacción de los factores biológicos y cognitivos. Los alumnos deben recordar que tienen que centrarse en el término de instrucción que se presenta.

Nivel de análisis cognitivo

Varios alumnos bien preparados escribieron respuestas muy buenas a esta pregunta. Sin embargo, algunos alumnos utilizaron estudios, como los de Phineas Gage y Broca, que no empleaban tecnología. Algunos alumnos describieron una amplia serie de investigaciones, pero no se centraron en un único proceso cognitivo, lo que les hizo recibir puntuaciones bajas.

La pregunta pedía que los alumnos discutieran el uso de la tecnología. Muchos alumnos evaluaron de manera automática la ética o la validez ecológica de los estudios, sin hacer referencia a alguna a la tecnología. Las mejores respuestas se centraron en la tecnología y en su aplicación.

Nivel de análisis sociocultural

Se han dado algunas respuestas muy buenas que demostraban amplitud de conocimientos y un pensamiento crítico bien desarrollado. Sin embargo, muchas respuestas eran demasiado descriptivas, sin pensamiento crítico añadido como idea adicional. Muchos ensayos aplicaron de forma repetida estrategias superficiales de evaluación como “como era un experimento de laboratorio, carecía de validez ecológica”. Un gran número de alumnos también incluyeron únicamente a Asch sin incorporar otros estudios, lo que dio como resultado puntuaciones bajas. Un número significativo de alumnos no discutió ningún punto fuerte de la investigación. Con frecuencia indicaron que los resultados eran los puntos fuertes. También hubo varias definiciones pobres de términos en las que los alumnos utilizaron los propios términos para definirlos: por ejemplo, conformidad es cuando uno está conforme. Se dieron muchos ejemplos en los que los alumnos no se centraron en la evaluación de la investigación, sino más bien en los tipos de conformidad o en los factores que influyen en la conformidad, lo que no era necesario. Por último, hubo varios alumnos que escribieron ensayos enteros sobre las técnicas de acatamiento o sobre la investigación sobre la obediencia.

Recomendaciones y orientación para la enseñanza a 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.
- También debe avisarse a los alumnos de que si la pregunta solo pide un estudio o teoría, solo se evaluará el primero. Los profesores deben aplicar este método de evaluación en sus propias prácticas con el fin de que los alumnos dejen el hábito de contar todo lo que saben, en lugar de centrarse en la pregunta.
- Cuando describan los estudios, los alumnos deben ser capaces de resumir de forma clara los objetivos generales, el procedimiento y los hallazgos, sin asumir demasiados conocimientos por parte del lector. Muchos alumnos escriben asumiendo que el lector conoce el estudio. También se da el caso de muchos alumnos que solo tienen una comprensión superficial de los estudios, con una comprensión limitada del procedimiento y los resultados.

- Es importante que se elija 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 la enseñanza no se centre en las investigaciones del siglo XIX. Además, es importante limitar la enseñanza de material que no sea pertinente para el currículo. Fueron varias las respuestas en el nivel cognitivo de análisis que se centraron en Freud, Pavlov, Skinner y Watson. También ha habido muchos alumnos que se centraron en Milgram al hablar de la conformidad. No está prohibido, por supuesto, enseñar estudios ajenos al currículo, pero los peores alumnos con frecuencia tienen dificultades para determinar lo que es pertinente en las preguntas del examen.
- Por último, es importante que durante la revisión para los exámenes, se oriente a los alumnos para que comprendan cómo aplicar los estudios para abordar los objetivos específicos de aprendizaje. Varios alumnos no utilizaron los estudios que eran pertinentes para el nivel de análisis enunciado en la pregunta.

Prueba 2 del Nivel Superior y del Nivel Medio

Bandas de calificación del componente

Nivel Superior

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 5	6 - 10	11 - 13	14 - 19	20 - 24	25 - 30	31 - 44

Nivel Medio

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 2	3 - 5	6 - 7	8 - 10	11 - 13	14 - 16	17 - 22

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

Como en exámenes anteriores, se han observado niveles de desempeño muy distintos en los colegios. Los alumnos de algunos colegios estaban muy bien preparados para hacer este examen, mientras que los alumnos de otros colegios no estaban preparados.

Por lo general, la mayoría de las respuestas mostraron cierto conocimiento de las ideas principales de las opciones, cierto conocimiento de la investigación adecuada, pero también un uso limitado del pensamiento crítico. Aunque muchos alumnos fueran capaces de proporcionar una evaluación rutinaria de los estudios, por ejemplo, falta de validez ecológica, pequeño tamaño de la muestra, problemas en la generalización y otros, hubo una escasez de análisis crítico del asunto principal de las preguntas. El término de instrucción “discutir” produjo varias respuestas mal organizadas en las que los alumnos escribieron todo lo que

sabían sobre el tema, aunque buena parte de la respuesta fuera tangencial al tema principal o sencillamente irrelevante. Los alumnos tuvieron especiales dificultades con el término de instrucción “contrastar” en este examen.

La mayor parte de los profesores indicaron que creían que el examen fue justo y de un nivel similar al examen del año pasado.

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

En su mayor parte, los alumnos fueron capaces de demostrar un conocimiento básico del contenido de las opciones y de la investigación adecuada. Las dos opciones más populares fueron la psicología del comportamiento anormal y la psicología de las relaciones humanas. Los alumnos de algunos colegios estaban muy bien preparados para responder a las preguntas de estas opciones. En las buenas respuestas, e incluso en un buen número de respuestas regulares, estuvo bien visible el conocimiento de los niveles de análisis. El nuevo programa de estudio está consiguiendo transmitir el tema fundamental de que la comprensión de la psicología humana requiere el examen de los factores biológicos, cognitivos y socioculturales.

Además de enseñar bien el contenido de las opciones, también es evidente que enseñar cómo escribir una respuesta al tipo de cuestiones utilizadas en la prueba 2 incide favorablemente en la calidad de los ensayos escritos por los alumnos. Los alumnos bien preparados con frecuencia comenzaron sus respuestas definiendo los términos clave, por ejemplo, “validez” o “fiabilidad”, y luego abordaron el meollo de la cuestión. El conocimiento de las exigencias concretas de los términos de instrucción permite a los alumnos escribir respuestas bien organizadas.

Puntos fuertes y débiles de los alumnos al abordar las distintas preguntas

Psicología del comportamiento anormal

Pregunta 1

Esta pregunta fue la más popular del examen. Las respuestas mostraron todas las variedades de calidad posibles, de las excelentes a las muy pobres. Las buenas respuestas comenzaron con definiciones claras de ambos términos. El uso acertado del estudio sobre los pseudo-pacientes de Rosenhan para discutir la validez formó parte casi siempre de las buenas respuestas. Los alumnos conocían mejor los estudios sobre la fiabilidad que los estudios sobre la validez. Las buenas respuestas incluyeron con frecuencia una discusión precisa de las investigaciones de Nicholls en el Great Ormond Street Children’s Hospital, de las investigaciones de Beck y otros y de Cooper y otros sobre la fiabilidad entre calificadores y de los estudios de Seeman sobre la fiabilidad test-retest del diagnóstico de esquizofrenia en las mujeres. La evaluación cultural y ética también estuvo con frecuencia bien integrada en las

mejores respuestas. Las mejores respuestas discutieron la forma cómo se interrelacionan la validez y la fiabilidad.

Las malas respuestas contenían definiciones vagas o inexactas de la validez y la fiabilidad, o simplemente “ponían juntos” ambos conceptos y no hacían ninguna distinción entre ellos. Los estudios de Rosenhan fueron entendidos mal por muchos alumnos. Muchos alumnos se limitaron a discutir los conceptos de normalidad y anormalidad o describieron los sistemas de clasificación en detalle, con poca o ninguna referencia a los temas de la validez y la fiabilidad, un enfoque que produjo muchas respuestas de poca calidad.

Pregunta 2

En comparación con la pregunta 1, muy pocos alumnos eligieron responder a esta pregunta. Los alumnos que sí contestaron a esta pregunta se centraron con frecuencia en la anorexia, la bulimia y las fobias. En general, las respuestas a esta pregunta fueron con frecuencia mediocres o pobres. Las buenas respuestas, de las que hubo pocas, evaluaron el enfoque de grupo tal y como se aplica concretamente en el tratamiento del trastorno identificado. Mucho más típicas fueron las respuestas que discutieron las características de la terapia de grupo de una forma general, que señalaron ventajas como recibir de apoyo de otras personas o mantener los costes bajos y desventajas como la falta de atención individual y la falta de confidencialidad. Con frecuencia se discutieron las terapias cognitivo-conductuales, aunque los alumnos no hicieron diferencias entre su uso en la terapia individual y en la terapia de grupo. Se ha visto un uso reducido de la investigación en la gran mayoría de las respuestas. Las malas respuestas también sustituyeron una discusión de los síntomas del trastorno elegido por una discusión que se centraba en el tratamiento en grupo del trastorno.

Unos pocos alumnos no identificaron de forma clara ningún trastorno en concreto y algunos alumnos eligieron la esquizofrenia de manera nada adecuada como el trastorno objeto de discusión en relación con el tratamiento de grupo.

Pregunta 3

En comparación con la pregunta 1, muy pocos alumnos eligieron responder a esta pregunta, pero hubo más respuestas a esta pregunta que a la pregunta 2.

Los alumnos bien preparados fueron capaces de proporcionar respuestas excelentes a esta pregunta. Las mejores respuestas discutieron múltiples factores biológicos. Los alumnos examinaron factores como las hormonas, los neurotransmisores, las estructuras cerebrales y la genética en lo que se refiere a la ansiedad y los trastornos alimentarios. Las buenas respuestas sobre las fobias incluyeron también una discusión de cómo la evolución puede preparar a los humanos para tener miedo de ciertos insectos, animales y situaciones. En general, los alumnos pudieron utilizar al menos alguna investigación en sus respuestas.

Algunos alumnos analizaron la depresión, identificándola como un trastorno de ansiedad, y unos pocos alumnos hablaron de la esquizofrenia, pero esas respuestas no les permitieron obtener ningún punto.

Psicología del desarrollo

Pregunta 4

En general, esta pregunta no se respondió muy bien. Algunos alumnos hicieron un trabajo razonable en la descripción de la fase de identidad de Erikson frente a la confusión de roles, pero tenían poco que decir sobre el segundo ejemplo de investigación. Las mejores respuestas compararon a Erikson con el estudio de la isla de Wight de Rutter.

Muy a menudo esta pregunta se contestó de una manera amplia y general proporcionando unos pocos estudios o teorías relativos a la adolescencia y proporcionando alguna evaluación de estas teorías. A menudo se ignoró el término de instrucción “contrastar”. Aunque se requerían dos estudios o teorías, en ocasiones solo se describió uno y el otro solo se indicó, lo que obviamente hizo que las respuestas fueran pobres.

Algunos alumnos interpretaron mal la pregunta. Se entendió que la pregunta requería dos estudios o teorías “que precedan a la adolescencia” en lugar de “en la adolescencia”. Algunos alumnos describieron las fases de desarrollo anterior a la adolescencia (dieron una visión general de las teorías de Piaget, Vygotsky o Freud), aunque con un gran grado de precisión, pero las respuestas eran totalmente irrelevantes para la pregunta que se pedía.

Las respuestas en español fueron pobres. Las respuestas fueron en su mayor parte relatos descriptivos de las teorías con una mínima atención a la adolescencia.

Pregunta 5

Pocos alumnos respondieron a esta pregunta. En el caso de aquellos que lo hicieron, la mayoría de las respuestas se centraron en definiciones y discusiones de la resiliencia, con poca atención a las estrategias para su desarrollo. Se dio una mejor comprensión de los factores de riesgo para los malos resultados de la vida que de las estrategias para desarrollar la resiliencia. El conocimiento y comprensión de los alumnos presentó limitaciones evidentes, y las estrategias se explicaban de manera superficial sin hacer referencia alguna a ningún estudio empírico.

Un número sustancial de respuestas ofreció respuestas anecdóticas y vagas en las que se resumían los programas de visitas a los hogares y los programas extraescolares, sin utilizar los estudios para la evaluación.

Pregunta 6

Esta fue una pregunta muy popular en esta opción. En conjunto, los alumnos mostraron una muy buena comprensión en sus respuestas y proporcionaron pruebas

válidas del pensamiento crítico relativas a los estudios o teorías que están relacionados con los posibles efectos de la privación o el trauma de la infancia en el desarrollo posterior. Las respuestas reflejaron los distintos enfoques que utilizaron los alumnos para abordar la pregunta, y la privación o el trauma se interpretaron de maneras diferentes. Se sacó a colación el estudio de caso de Genie, junto con descripciones de niños “salvajes”, aunque a menudo se incluyó también la privación económica. Muchos alumnos analizaron también el estudio longitudinal de los gemelos checos de Koluchova. Algunos alumnos utilizaron los estudios de Harlow sobre la privación en los monos rhesus con muy buenos resultados. Algunas de las buenas respuestas también discutieron las investigaciones de Hazen y Shaver sobre el apego en la infancia y su efecto en las relaciones románticas adultas. También se presentaron el concepto de Bowlby de un modelo de trabajo interno y el estudio de Rutter sobre niños adoptados rumanos, a menudo con discusiones bien preparadas sobre la resiliencia.

Muchas de las peores respuestas indicaron y discutieron las investigaciones psicológicas relacionadas con la privación o el trauma, pero no las desarrollaron posteriormente, o los alumnos discutieron los efectos de la privación o el trauma, sin proporcionar suficientes estudios para respaldar la discusión.

Psicología de la salud

Pregunta 7

Esta pregunta fue la más popular de esta opción. La mayor parte de los alumnos fueron capaces de resumir dos estresores. Muchos de los alumnos optaron por identificar dos tipos de estresores –agudos y crónicos–, y en algunas ocasiones proporcionaron también ejemplos de cada tipo. Muchos alumnos también ofrecieron detalles de estos estresores, o entraron en una discusión completa del modelo del síndrome general de adaptación de Selye, que la pregunta no requería.

En general, se ha visto un buen conocimiento descriptivo de las estrategias de afrontamiento, aunque escaseaban las pruebas de pensamiento crítico.

Unos pocos alumnos no pudieron identificar correctamente una estrategia para afrontar el estrés. Sin embargo, el mayor problema para responder a esta pregunta fue que algunos alumnos emplearon demasiado tiempo y esfuerzo para responder a la primera parte de la pregunta, lo que ocasionó evaluaciones demasiado breves de una estrategia para afrontar el estrés.

Pregunta 8

La mayoría de las respuestas a esta pregunta fueron pobres. La mayor parte de las respuestas ofrecieron un conocimiento y comprensión superficiales y no discutieron e identificaron con precisión una estrategia de prevención. Fueron muchas las respuestas de sentido común que se ofrecieron a esta pregunta. Muchas respuestas abordaron las formas de reducir la obesidad, en lugar de cómo prevenirla.

Pregunta 9

La mayor parte de los alumnos se centraron en los tratamientos para los comportamientos adictivos, la mayor parte de las veces la adicción al alcohol o al tabaco. Las respuestas fueron con frecuencia descriptivas y genéricas. Los alumnos proporcionaron a menudo una revisión de un tratamiento médico o una terapia cognitivo-conductual. Normalmente, el tratamiento médico contaba con un conocimiento más detallado y una discusión en profundidad, mientras que el conocimiento de la terapia cognitivo-conductual con frecuencia era superficial y menos detallado. Las respuestas fueron con frecuencia especialmente superficiales cuando los alumnos eligieron los grupos de apoyo social o las terapias familiares como modo de tratamiento.

Nota adicional: Algunos examinadores señalaron una manifiesta debilidad en la formulación de las preguntas de la opción de psicología de la salud. La debilidad se observó sobre todo en la falta de conocimientos basados en investigaciones, unido a la incapacidad de contestar de manera coherente y adecuada a las exigencias de los términos de instrucción en el nivel superior del pensamiento crítico (por ejemplo, la evaluación no se completó de forma coherente).

Psicología de las relaciones humanas**Pregunta 10**

Esta pregunta fue popular en esta opción. La mayor parte de los alumnos fueron capaces de proporcionar una discusión precisa y completa, estableciendo una conexión entre los factores biológicos y su efecto en las relaciones humanas.

Se abordaron varios factores biológicos diferentes: los neurotransmisores, las hormonas, los centros de placer del cerebro, los centros de agresión del cerebro, las teorías evolucionistas aplicadas a la selección natural de los rasgos de supervivencia y la selección sexual aplicada a la selección de compañeros. Las mejores respuestas abordaron directamente el término de instrucción “en qué medida” e indicaron las limitaciones e incluso el reduccionismo de limitarse a centrarse en los factores biológicos, e incluyeron algún tipo de discusión del papel de las influencias cognitivas y socioculturales en las relaciones humanas.

Algunos alumnos interpretaron el término “relaciones humanas” como “relaciones

interpersonales” y, por tanto, se centraron únicamente en la atracción y la selección de compañeros, aunque aún así consiguieron elaborar buenas respuestas. Otros alumnos se centraron en las tres partes de la opción: responsabilidad social, relaciones interpersonales y violencia. Este enfoque también produjo algunas respuestas muy buenas. Era suficiente utilizar la prueba empírica dada para contestar bien la pregunta.

Unas pocas respuestas mostraron un conocimiento limitado de la conexión entre los factores biológicos y las relaciones humanas. En lugar de centrarse en las relaciones humanas, estas respuestas proporcionaron una breve explicación de los factores biológicos que afectan al comportamiento humano en general.

Pregunta 11

Esta pregunta recibió un conjunto muy variado de respuestas, incluidas algunas respuestas excelentes y otras que eran meramente respuestas de sentido común. El conocimiento descriptivo de la investigación fue mucho mejor que la capacidad de aplicar la evaluación y el análisis críticos para investigar sobre la pregunta. Muchos alumnos recurrieron a indicar que era muy complicado el tema de por qué las relaciones cambian o terminan, y esa afirmación reemplazó cualquier análisis con una mayor profundidad. Las mejores respuestas discutieron con frecuencia la teoría del cambio social, la teoría de la equidad y la teoría de la atribución. Algunas de las mejores respuestas integraron los factores biológicos en la discusión.

Esta pregunta, así como la pregunta 14 sobre psicología del deporte, parecen haber sido las preguntas “comodín” de este examen. Si un alumno no sabía contestar a la segunda pregunta, esta pregunta parecía ofrecer una posibilidad razonable de conseguir algún punto para algunos alumnos. Los alumnos que utilizaron esta estrategia para seleccionar la pregunta a contestar elaboraron respuestas muy pobres basadas en un saber popular no demostrado que carecía de pruebas procedentes de la investigación psicológica.

Pregunta 12

Esta pregunta fue una elección muy popular en esta opción. En muchos casos, los examinadores señalaron que las respuestas a esta pregunta recibieron puntuaciones altas. Muchos alumnos escribieron respuestas bien desarrolladas y bien organizadas, respaldadas por teorías apropiadas, estudios y evaluación de la investigación utilizada. Algunas de las mejores respuestas señalaron la incidencia mucho más alta del efecto espectador en los experimentos de laboratorio de Latané y Darley que en el experimento de campo del metro de Nueva York de Piliavin y otros. Las respuestas en las notas medias presentaron una evaluación limitada. Por ejemplo, las investigaciones de Latané y Darley que estudiaron el papel del número de personas disponibles para ayudar (difusión de la responsabilidad), así como la influencia social informativa (ignorancia pluralista), se describió a menudo, pero no se evaluó de forma clara. Los factores culturales fueron introducidos como comentarios de evaluación. Por lo general, las respuestas que analizaron un número relativamente menor de teorías o estudios sobre los factores que influyen en el

efecto espectador con mayor profundidad a menudo obtuvieron más puntos, pero algunas respuestas muy sólidas mostraron una amplia variedad de conocimientos gracias a la descripción y evaluación de las investigaciones de Piliavin y Bateson y Darley, además del estudio de Latané y Darley utilizado con mayor frecuencia.

Algunos alumnos utilizaron teorías sobre el altruismo con buenos resultados, pero otros utilizaron esas teorías sin vincularlas al efecto espectador. Algunos alumnos emplearon demasiado tiempo describiendo el caso de Kitty Genovese con profusión de detalles (en ocasiones inexactos).

Psicología del deporte

Pregunta 13

Hubo muchas respuestas a esta pregunta. Muchos alumnos fueron capaces de hablar de investigaciones pertinentes para la fijación de metas, aunque la evaluación de la investigación fuera muy limitada. Una respuesta corriente se limitó a discutir la fijación de metas en relación con el desempeño. Otro problema habitual que señalaron los examinadores fue que los alumnos proporcionaban respuestas descriptivas que hablaban de la motivación intrínseca y extrínseca y de poco más

Pregunta 14

Los alumnos proporcionaron una serie de respuestas diferentes a esta pregunta. En algunos casos, esta pregunta generó respuestas pertinentes y bien formuladas. Sin embargo, la mayoría de las respuestas enumeraron las razones para utilizar drogas en el deporte, pero no respaldaron sus conocimientos con pruebas empíricas. De manera similar a la opción de la psicología de las relaciones humanas de la pregunta 11, algunos alumnos parece que escogieron esta pregunta cuando veían que no sabían contestar una segunda pregunta. Varias de las respuestas no contenían más que conocimientos generales, por ejemplo, los atletas usan los esteroides para aumentar su potencia.

Pregunta 15

Hubo muy pocas respuestas a esta pregunta. Los alumnos que respondieron a la pregunta proporcionaron respuestas no muy buenas con poco o ningún uso de la investigación. Se manejaron de manera muy pobre las exigencias del término de instrucción “en qué medida”.

Recomendaciones para la enseñanza a futuros alumnos

- Parece que el principal problema para los alumnos es la incapacidad de interpretar adecuadamente los términos de instrucción. Por tanto, desde el mismo comienzo del curso, los alumnos deben estar familiarizados con los términos de instrucción y se les debe enseñar claramente las diferentes exigencias que recaen en el lector. Debe hacerse a los alumnos los mismos tipos de preguntas que se les hará luego en los exámenes, para que estén bien preparados para los exámenes finales del IB.

- Debe darse a los alumnos las preguntas de los exámenes anteriores para que las hagan y, una vez que las hayan completado, el profesor deberá explicar los criterios y los esquemas de calificación a los alumnos. Una vez que se haya hecho esta revisión, debe pedirse a los alumnos que evalúen su propio trabajo. De esta forma, los alumnos pueden responsabilizarse de su propio aprendizaje.
- Fundamentalmente, los profesores deben asegurarse de que los alumnos establezcan una conexión entre la teoría, el concepto o el término y los estudios empíricos, y que se aseguren cuando lo hagan de que están evaluando el concepto, no únicamente los estudios empíricos, de acuerdo con los requisitos de la pregunta.
- Debe recordarse a los alumnos que añadir varios estudios cuando se pide un estudio rebaja la calidad de la respuesta, ya que solo se tendrá en cuenta el primero. Si una pregunta pide una discusión o evaluación de un estudio o teoría, se pueden introducir estudios o teorías adicionales como parte de la discusión, que pueden hacer ganar puntos solo si la investigación se utiliza clara y explícitamente para evaluar la primera investigación que se discute.
- Los profesores no deben animar a los alumnos a proporcionar un alto número de estudios, ya que con frecuencia los recuerdan mal y no conservan la pertinencia necesaria con la pregunta. En su lugar, deben centrarse en un estudio clásico y en uno o dos ejemplos actualizados, y luego en aplicar esa investigación para contestar adecuadamente la pregunta.
- En los colegios que son nuevos en el Programa del Diploma del IB, es fundamental que los profesores sean conscientes de las exigencias del programa de Psicología. Las discusiones genéricas y anecdóticas no son respuestas aceptables a las preguntas de los exámenes. Deben enseñarse terminología y conceptos psicológicos concretos. Además, debe enseñarse a los alumnos a desglosar las exigencias de las preguntas y a estructurar y respaldar adecuadamente las respuestas.

Prueba 3 del Nivel Superior

Bandas de calificación del componente

Calificación final:	1	2	3	4	5	6	7
Puntuaciones:	0 - 1	2 - 3	4 - 7	8 - 10	11 - 13	14 - 16	17 - 30

Comentario general

En conjunto, los alumnos parece que han entendido bien el material de estímulo este año, pero algunos alumnos han tenido problemas para integrar el material de estímulo con conocimiento y comprensión de la metodología cualitativa de investigación.

El propósito de la prueba 3 es demostrar el conocimiento de los métodos cualitativos de investigación, tal y como se resume en la guía. Fueron bastantes los alumnos que parecían no estar preparados para la prueba 3 y mostraron poco o ningún conocimiento y comprensión de los métodos cualitativos de investigación. A menudo se ha visto que los razonamientos se basaban en el conocimiento de la investigación cuantitativa y, en concreto, en la investigación experimental. Al igual que en años anteriores, la mayoría de los alumnos utilizaron el término “experimento” como sinónimo de “estudio de investigación”, lo que no es recomendable.

Algunos alumnos tuvieron problemas para utilizar adecuadamente la información del material de estímulo. Por ejemplo, en lugar de utilizar el estudio de caso con cinco participantes del material de estímulo como punto de inicio para la explicación y la discusión, los alumnos sugirieron varios métodos de muestreo, incluido el muestreo aleatorio, que podían aumentar el número de participantes o sugirieron cambiar el estudio.

Otro problema que se ha observado este año fue la tendencia a centrarse demasiado en el desastre y el estrés, y menos en el estudio de caso y en la metodología de investigación cualitativa. Por ejemplo, los alumnos sugerían formas de abordar el trabajo de ayuda, en lugar de centrarse en lo que estaban estudiando los investigadores en el estudio de caso. Además, algunos alumnos se centraron demasiado en el estrés y en cómo podían ayudar los dos psicólogos a los trabajadores de emergencias, en lugar de responder a las preguntas sobre los métodos cualitativos de investigación. En conjunto, en algunos casos hubo dificultades para integrar el material de estímulo con las preguntas, lo que produjo en muchos casos que no se centraran bien en la pregunta.

Hubo cierta dispersión en las calificaciones que se otorgaron y los alumnos obtuvieron puntuaciones de todo tipo, algunos en las puntuaciones más bajas, la mayoría en las puntuaciones medias y unos pocos en las puntuaciones altas. Esto indica que muchos alumnos no estaban bien preparados para contestar a las preguntas de la prueba 3.

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

Parece que la pregunta más difícil para los alumnos fue la pregunta 2, que incluía el conocimiento y comprensión hasta el punto que sea posible para generalizar a partir de un único estudio de caso. Muchos alumnos utilizaron los conocimientos relativos a la investigación cuantitativa (generalización a partir de una población) y los alumnos que obtuvieron las puntuaciones más altas fueron capaces de compararla con la generalización a partir de un único estudio de caso utilizando los conocimientos pertinentes. El término de instrucción “en qué medida” con frecuencia no se abordó de manera adecuada. Bastantes de los alumnos demostraron conocer la generalización inferencial, representativa o teórica y la relacionaron con el material de estímulo en la respuesta, aunque el conocimiento de esos términos no fuera necesario para obtener las mejores puntuaciones.

Por lo que se refiere a la pregunta 1, parece que algunos alumnos no sabían que el muestreo aleatorio no es un método de muestreo apropiado para un estudio de caso con pocos participantes. Sin embargo, la mayoría de los alumnos fueron capaces de indicar un método

de muestreo pertinente y de explicarlo en relación con el material de estímulo. Algunos alumnos explicaron que el muestreo de conveniencia se utiliza en el estudio en el material de estímulo y luego sugirieron métodos alternativos, como el muestreo intencionado, y explicaron por qué este método también podría ser apropiado en el estudio contenido en el material de estímulo. Otros sugirieron el muestreo de bola de nieve como el único método de muestreo o como una posibilidad si se ampliara el estudio. En la pregunta 3, muchos alumnos tuvieron dificultades para hablar de cómo puede utilizarse un estudio de caso para investigar un problema y solo escribieron sobre los métodos de obtención de datos y sobre cuestiones éticas.

En algunos exámenes con frecuencia se basó el análisis en la especulación, en lugar de en el conocimiento de la investigación cualitativa aplicada al material de estímulo.

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

Muchos alumnos mostraron una buena comprensión de los métodos de muestreo en la pregunta 1 y de los métodos de obtención de datos en la pregunta 3. Por lo que se refiere a la generalización de un solo estudio de caso, algunos alumnos parecían estar bien preparados y mostraron su comprensión de los principios de la generalización en la investigación cualitativa y fueron capaces de explicar cuándo y cómo puede hacerse, a menudo haciendo referencia a términos teóricos, lo que indica un sólido conocimiento de las cuestiones de la generalización en un solo estudio de caso.

Puntos fuertes y débiles de los alumnos al abordar las distintas preguntas

Pregunta 1

La pregunta fue bien contestada por lo general, aunque unos cuantos alumnos escribieron que el muestreo aleatorio sería el más adecuado. El muestreo aleatorio no es apropiado en un estudio de caso o en una investigación cualitativa. Respuestas como esas argumentaron en la mayoría de los casos que una muestra debe ser representativa de la población objetivo y que eran necesarios muchos participantes para lograrlo. Esas respuestas apenas hicieron referencia al material de estímulo o pasaron por alto la información contenida en el material de estímulo que decía que la muestra constaba únicamente de cinco participantes.

La mayoría de los alumnos fueron capaces de indicar un método de muestreo pertinente y las mejores respuestas explicaron también por qué era apropiado el método de muestreo en este estudio de caso, además de que métodos alternativos podían utilizarse y por qué, utilizando la información contenida en el material de estímulo como apoyo. Algunos alumnos explicaron que el muestreo de conveniencia se utiliza en el estudio contenido en el material de estímulo y sugirieron con frecuencia métodos alternativos, como el muestreo intencionado, y explicaron por qué este método también podría ser apropiado en el estudio contenido en el material de estímulo. Sin embargo, otros sugirieron el muestreo de bola de nieve como el

único método de muestreo o como una posibilidad si se ampliara el estudio.

Pregunta 2

Parece que la pregunta fue especialmente difícil para varios alumnos que consiguieron cero o muy pocos puntos en esta pregunta. Las respuestas que alcanzaron puntuaciones bajas generalmente hicieron referencia al tipo de desastre como un criterio para la generalización, o a la cultura, el sexo y la edad como determinantes del grado de generalización que podía hacerse, lo que muestra un conocimiento y comprensión limitados de la generalización en la investigación cualitativa o a partir de un solo estudio de caso. Sin embargo, parece que algunos alumnos estaban muy bien preparados y fueron capaces de discutir la generalización a partir de un solo estudio de caso que se basaba en una gran abundancia de datos y hallazgos de otros estudios de casos similares. Muchos alumnos estaban también familiarizados con la terminología relativa a la generalización en la investigación cualitativa, es decir, la generalización inferencial, representativa o teórica, y las respuestas que alcanzaron las puntuaciones más altas también fueron capaces de explicar esto en relación con el material de estímulo.

Pregunta 3

Las peores respuestas no abordaron cómo puede utilizarse un estudio de caso para investigar un problema ni sugirieron métodos pertinentes de obtención de datos. Las peores respuestas no abordaron el uso de un estudio de caso y se limitaron a describir uno o dos métodos de obtención de datos de un modo bastante genérico, apenas utilizando el material de estímulo. En algunas respuestas se afirmó que la observación de los participantes, e incluso la observación encubierta de los no participantes, sería apropiada para el estudio, aunque eso sería difícil si se utilizara de forma apropiada la información del material de estímulo, ya que se mencionaba expresamente que se había obtenido el consentimiento informado y que los investigadores acompañaron al equipo elegido a la zona del desastre natural.

Las mejores respuestas explicaron qué es un estudio de caso y cómo puede utilizarse para investigar un problema como el del material de estímulo y luego explicaron dos o tres métodos pertinentes de obtención de datos y por qué serían especialmente pertinentes para este estudio de caso. Esas respuestas normalmente se centraron en la observación de los no participantes y en varias formas de entrevistas como métodos pertinentes de obtención de datos. Por ejemplo, muchos explicaron que las entrevistas a los grupos focales podían ser un método útil, ya que los participantes trabajaban como equipo y la idea de la investigación era averiguar cómo se ayudaban entre sí los miembros de un equipo en entornos estresantes.

Recomendaciones para la enseñanza a futuros alumnos

- Parece que el principal reto en la prueba 3 es que los alumnos deben aprender a distinguir entre los métodos de investigación cuantitativa y cualitativa. Es importante que los alumnos utilicen sobre todos sus conocimientos sobre los métodos de investigación cualitativa cuando respondan a la prueba 3. Un segundo reto es

aprender a utilizar el material de estímulo de una forma adecuada para respaldar el argumento. La prueba 3 del programa de estudios actual requiere que los alumnos integren el conocimiento de los métodos cualitativos de investigación en el material de estímulo. No es suficiente describir lo que hay en el material de estímulo o utilizarlo para distintas especulaciones sin hacer referencia a la metodología. El material de estímulo está destinado a servir como punto de inicio para el análisis de cómo pueden aplicarse los métodos de investigación cualitativa a un estudio concreto.

- Los alumnos deben practicar “lo que es ser como un investigador cualitativo” para que sean capaces de aplicar el conocimiento metodológico pertinente al material de estímulo. Esto significa que la enseñanza debe incluir la exposición a varios estudios cualitativos y preferentemente también dar a los alumnos la oportunidad de llevar a cabo algún tipo de proyecto de investigación menor, con el fin de comprender mejor el razonamiento de un investigador cualitativo. También es importante que los alumnos lleguen a entender la diferencia entre los métodos cuantitativos y cualitativos para que eviten utilizar demasiada terminología de los métodos cuantitativos. Por ejemplo, muchos alumnos no parecen ser conscientes de que términos como “investigador”, “experimento” y “efecto Hawthorne” no son realmente apropiados para su uso en el contexto de un estudio cualitativo.
- Las preguntas que se basan en una introducción, sin un resumen de un estudio o situación (el material de estímulo), deben proporcionar a los alumnos una oportunidad para demostrar cómo aplicar el conocimiento y comprensión pertinentes de los métodos de investigación cualitativa en el contexto de ese estudio o situación. En el material de estímulo se han numerado las líneas de cinco en cinco, de forma que los alumnos puedan hacer referencia a las líneas sin tener que utilizar extensas citas. Esto puede utilizarse con mayor eficacia en las respuestas y la enseñanza de esta prueba debe incluir mostrar a los alumnos cómo encontrar las partes pertinentes del material de estímulo que puedan respaldar la explicación de la metodología de investigación.
- Esto significa que los alumnos deben estar preparados de tal forma que posean un conocimiento general de los métodos de investigación cualitativa, tal y como se resumen en la guía, y una destreza en la aplicación de estos conocimientos en lo que se refiere al material de estímulo. La utilización de preguntas de exámenes anteriores podría facilitar esto o los profesores podrían simplemente preparar ellos mismos el material y las preguntas cuando impartan los métodos de investigación cualitativa. Se recomienda también que practiquen con los alumnos la elaboración de evaluaciones y discusiones equilibradas, en lugar de afirmaciones y especulaciones sin soporte alguno.
- Se aconseja también a los profesores que enseñen a los alumnos lo que significa responder a los términos de instrucción relativos a la prueba 3, por ejemplo lo que significa “explicar”. Son demasiados los alumnos que se limitan a especular sobre lo que piensan que es pertinente, pero que no explican por qué podría ser pertinente en relación con el material de estímulo.