

Estado actual de las investigaciones sobre validación de exámenes de lengua extranjera

Lorena Llosa



New York University

18 de agosto del 2006

Objetivos

- El argumento del uso de la evaluación (Bachman, 2005)
- Las cualidades de utilidad
 - Confiabilidad
 - Validez de constructo
 - Autenticidad
 - Impacto
- El uso del argumento del use de la evaluación para el diseño de examen para las clases de lengua

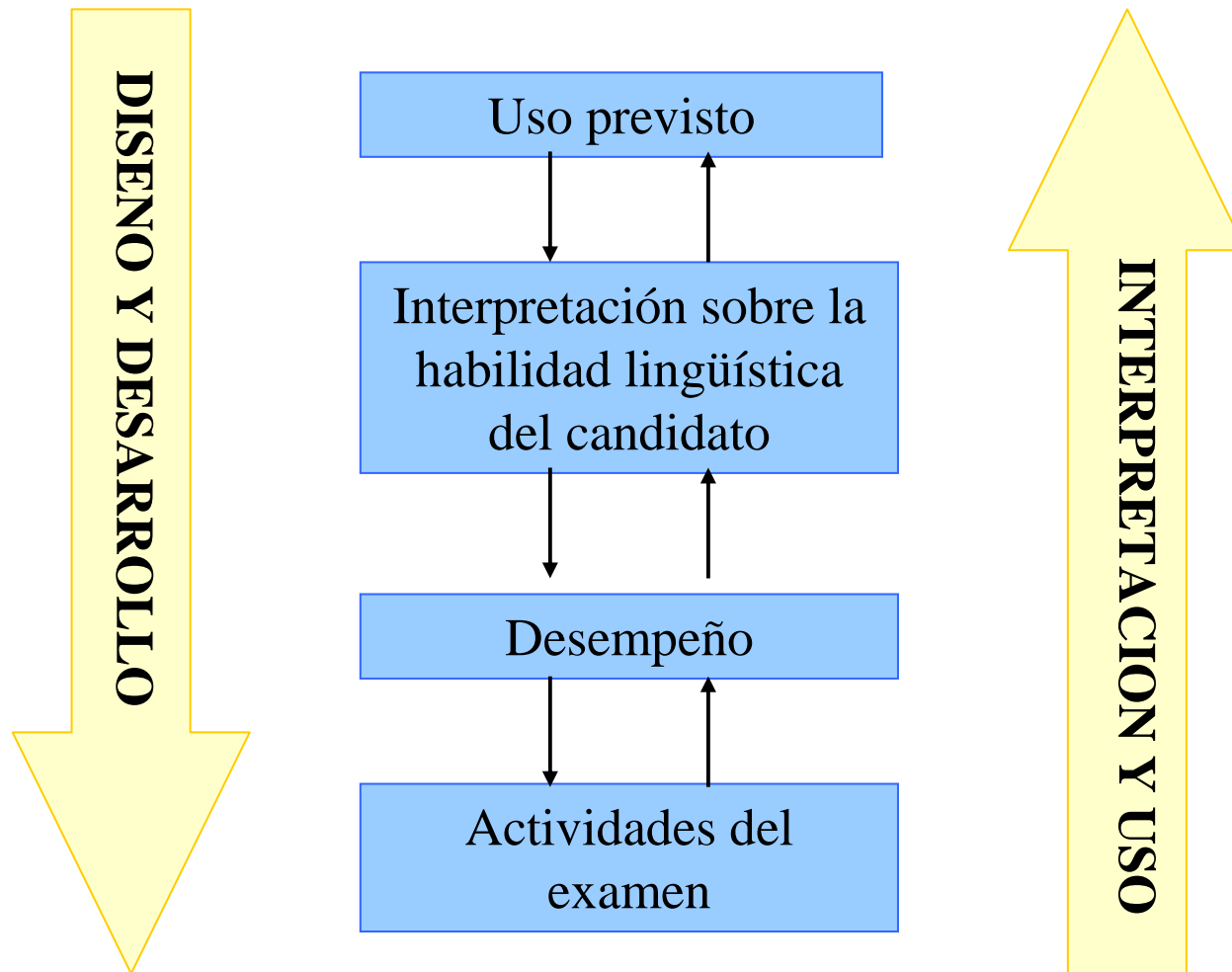
Usos de evaluaciones de lengua

- Tomar decisiones que afectan a individuos
 - Seleccionar estudiantes
 - Asignarlos a un nivel adecuado
 - Evaluar su progreso y logro
 - Dar notas
 - Acreditar su dominio
- Tomar decisiones que afectan a programas, instituciones
- Para investigación

Usamos evaluaciones de lengua para tomar estas decisiones porque consideramos que el individuo necesita determinada habilidad lingüística para desempeñar con éxito ciertas actividades fuera del marco del examen.

Por lo tanto, cuando tomamos una evaluación de lengua, nuestra intención es interpretar el desempeño en este examen, reflejado en un resultado, como un indicador de algún aspecto de la habilidad lingüística del individuo.

De desempeño a uso de una evaluación



La consideración más importante al diseñar y desarrollar un examen de lengua es el **uso** (“*use*”) que tenemos en mente, y por eso la cualidad más importante de un examen es su **utilidad** (“*usefulness*”)

Para demostrar la utilidad de un determinado examen para el propósito previsto y asumir la responsabilidad ante las partes interesadas, tenemos que

- (1) formular una base lógica que justifique los usos y decisiones y
- (2) proporcionar evidencia que dé apoyo a esa base lógica.

Argumento de uso de evaluación (Bachman, 2005)



El argumento del uso de la evaluación (“*assessment use argument*”), nos proporciona el fundamento y la justificación de las decisiones que tomamos cuando diseñamos y desarrollamos un examen.

También nos provee un marco teórico lógico necesario para justificar las interpretaciones y el uso que hacemos basándonos en el desempeño del candidato.

¿Por qué necesitamos un argumento del uso de la evaluación?

Las partes interesadas nos pueden pedir que proporcionemos una justificación, basada en evidencia, de que la evaluación es útil para el propósito previsto.

El argumento del uso de la evaluación nos proporciona el marco teórico para cumplir este objetivo.

¿A quiénes tenemos que convencer?

- Nosotros mismos
- Los candidatos (nuestros alumnos)
- Otros profesores
- Otras partes interesadas

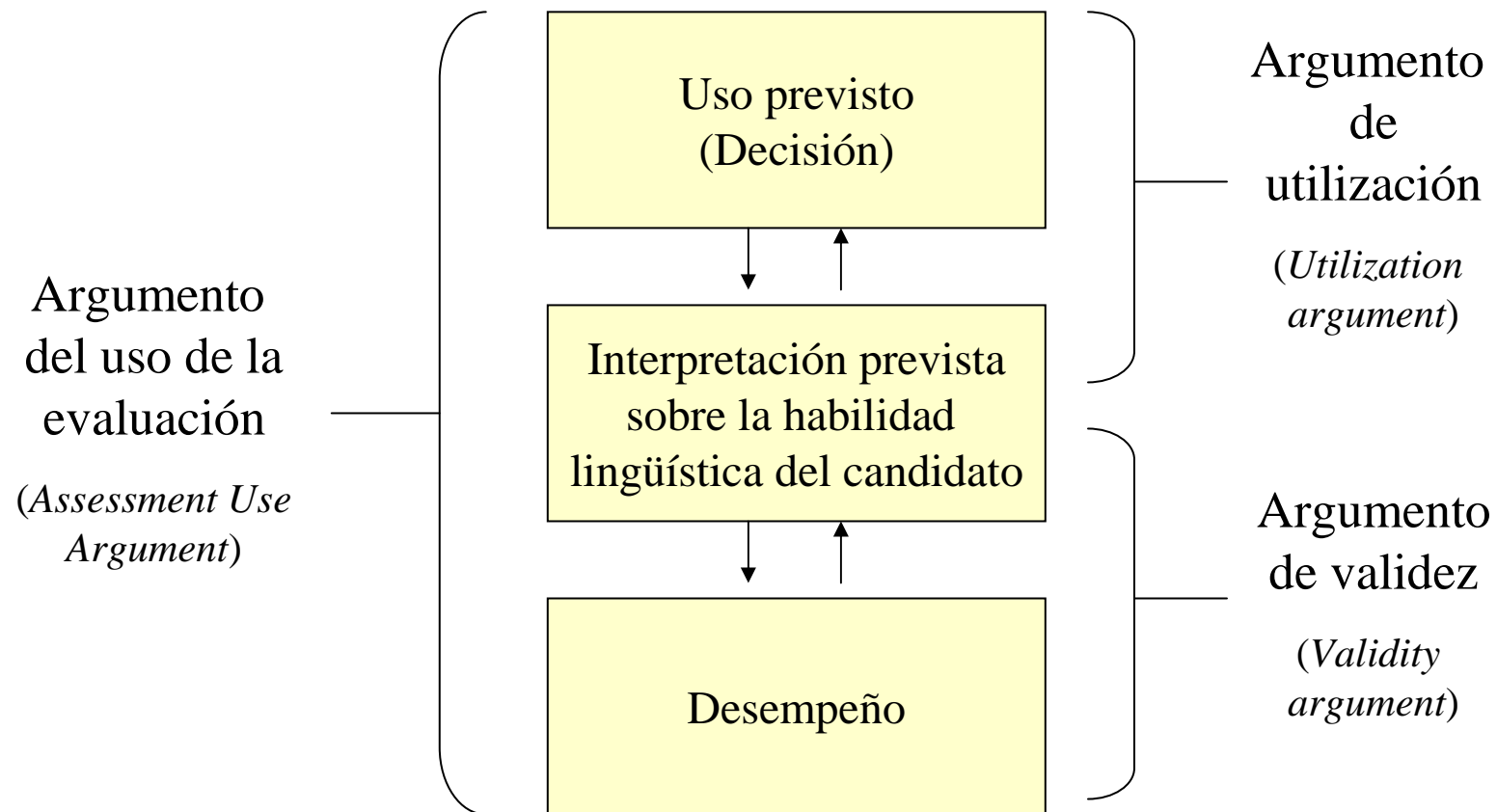
Validación

El proceso que seguimos para evaluar la utilidad de un determinado examen se llama **validación** e incluye dos actividades interrelacionadas:

1) formular un argumento del uso de la evaluación

2) recopilar evidencia pertinente para apoyar las interpretaciones y los usos previstos de los resultados del examen

Argumento del uso de la evaluación



Elementos del argumento del uso de la evaluación (Toulmin, 2003)

DECLARACIONES (“*claims*”): afirmaciones respecto de las interpretaciones y de los usos que queremos darle al desempeño en el examen.

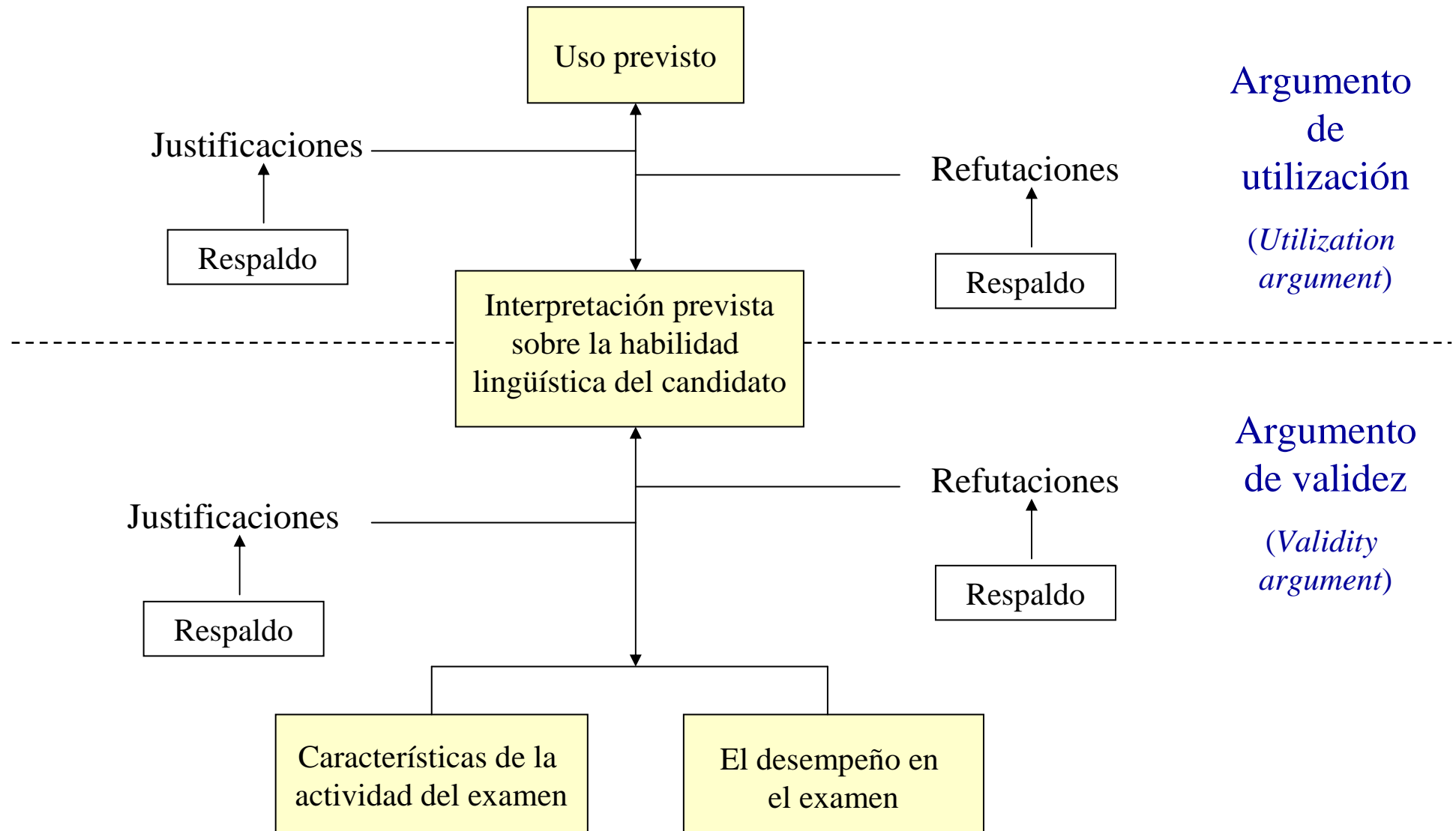
JUSTIFICACIONES (“*warrants*”): afirmaciones que justifican las interpretaciones y los usos. Incluyen afirmaciones sobre varias cualidades del examen que aportan a su utilidad total.

REFUTACIONES (“*rebuttals*”): afirmaciones respecto de otras interpretaciones posibles del desempeño del candidato, posibles razones de no usar el examen, o posibles consecuencias no planeadas de usar el examen de la manera prevista.

RESPALDO (“*backing*”): la evidencia recopilada para dar apoyo a las justificaciones y para rechazar o debilitar las refutaciones.

DATOS (“*data*”): el desempeño del candidato y las características de las actividades del examen.

Elementos de un argumento del uso de la evaluación



La utilidad de la evaluación (“*Test usefulness*”)

- Las cualidades de utilidad (“*the qualities of usefulness*”)
 - Confiabilidad
 - Validez de constructo
 - Autenticidad
 - Impacto

Confiabilidad

- Consistencia en la medida de un determinado constructo o habilidad
- Fuentes potenciales de errores de medición:
 - El contexto
 - El proceso de administración
 - Los candidatos
 - Las características de las actividades del examen
 - El modo de calificación
- Nuestro objetivo es minimizar los errores de medición que están bajo nuestro control por la manera en que diseñamos las evaluaciones.

Validez de constructo

La significatividad y la propiedad de las interpretaciones que hacemos a partir del desempeño en el examen.

Se refiere a la medida en que podemos interpretar los resultados de un determinado examen como indicadores de las habilidades o constructos que queremos evaluar.

Un constructo es una habilidad o atributo que definimos de manera específica y sirve de base para el diseño de un examen o actividad de examen y para interpretar los resultados obtenidos.

Confiabilidad y validez de constructo

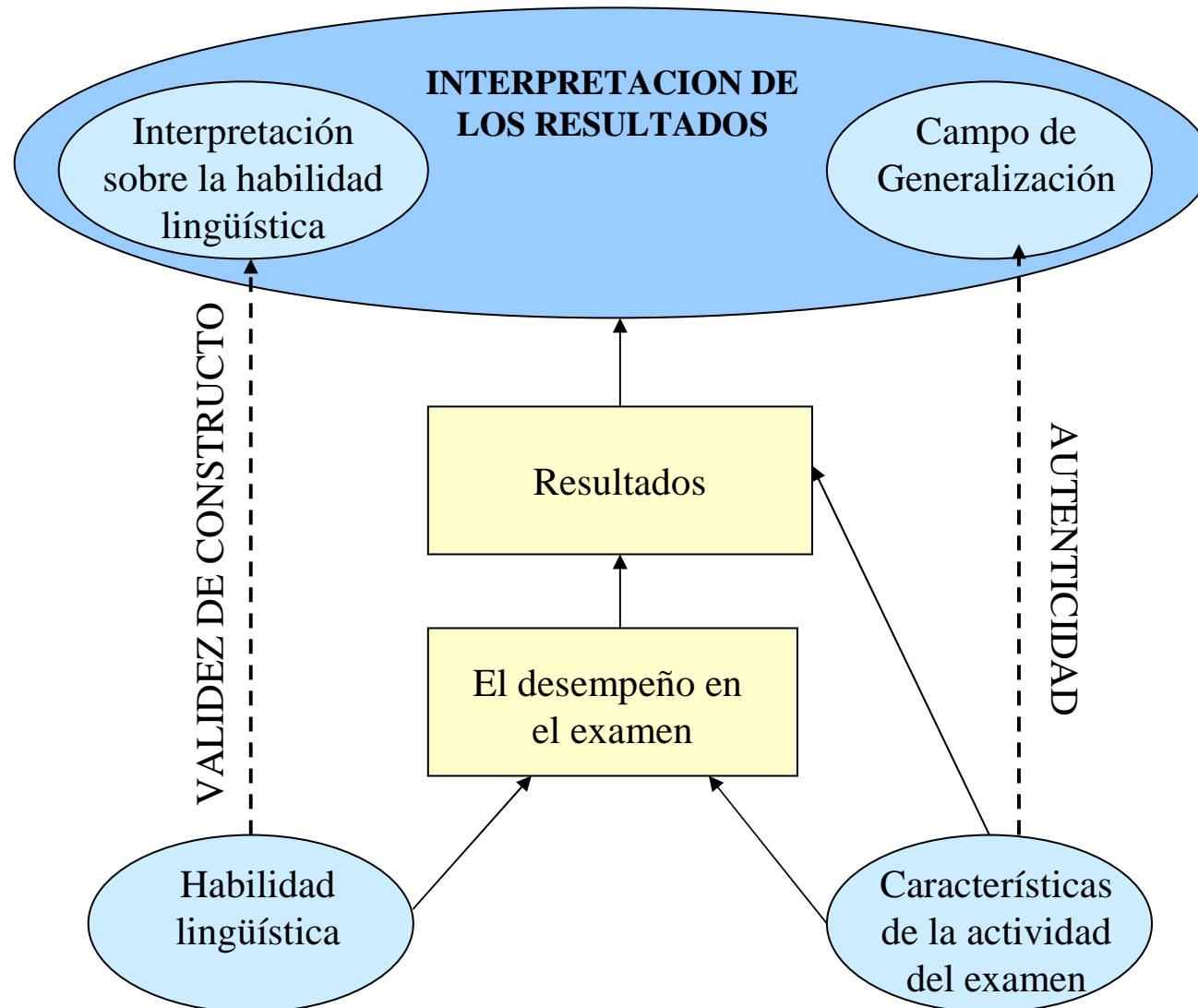
La confiabilidad es una cualidad de los **resultados** del examen mientras que la validez de constructo es una cualidad de las **interpretaciones** que hacemos a partir de los resultados.

La confiabilidad es una condición necesaria para la validez de constructo y la utilidad total del examen. Pero, la confiabilidad no es una condición suficiente ni para la validez de constructo ni para la utilidad total del examen.

Autenticidad

- Definimos la autenticidad como el grado o nivel de correspondencia entre las características de una determinada actividad en el examen y las características de una actividad en el campo donde se usa la lengua meta.
- La autenticidad tiene que ver con la medida en que las interpretaciones basadas en el desempeño en la evaluación pueden ser generalizadas mas allá del contexto del examen.

La validez de constructo y la autenticidad



Autenticidad

- Definimos la autenticidad como el grado o nivel de correspondencia entre las características de una determinada actividad en el examen y las características de una actividad en el campo del uso de la lengua meta.
- La autenticidad tiene que ver con la medida en que las interpretaciones basadas en el desempeño en la evaluación pueden ser generalizadas mas allá del contexto del examen.
- También se refiere a la manera en que los candidatos perciben a las actividades del examen.
- Finalmente, la autenticidad es relativa. Debemos hablar de actividades que son relativamente más o menos auténticas en vez de decir que una actividad es auténtica o inauténtica.

Impacto

- El mero acto de tomar y rendir un examen implica ciertos valores y objetivos y tiene consecuencias. De la misma manera, la forma en que usamos los resultados también supone valores y objetivos y esos usos también tienen consecuencias.
- El impacto puede ocurrir en dos niveles: a nivel de los individuos y a nivel de sistemas educativos o la sociedad

Impacto sobre los individuos

- Los candidatos pueden ser afectados por tres aspectos del proceso de la evaluación:
 - la experiencia de rendir y prepararse para el examen
 - las notas, comentarios y observaciones que reciben sobre su desempeño en el examen
 - las decisiones a tomar a partir de su desempeño
- Los usuarios
 - Profesores
 - Directores
 - Padres

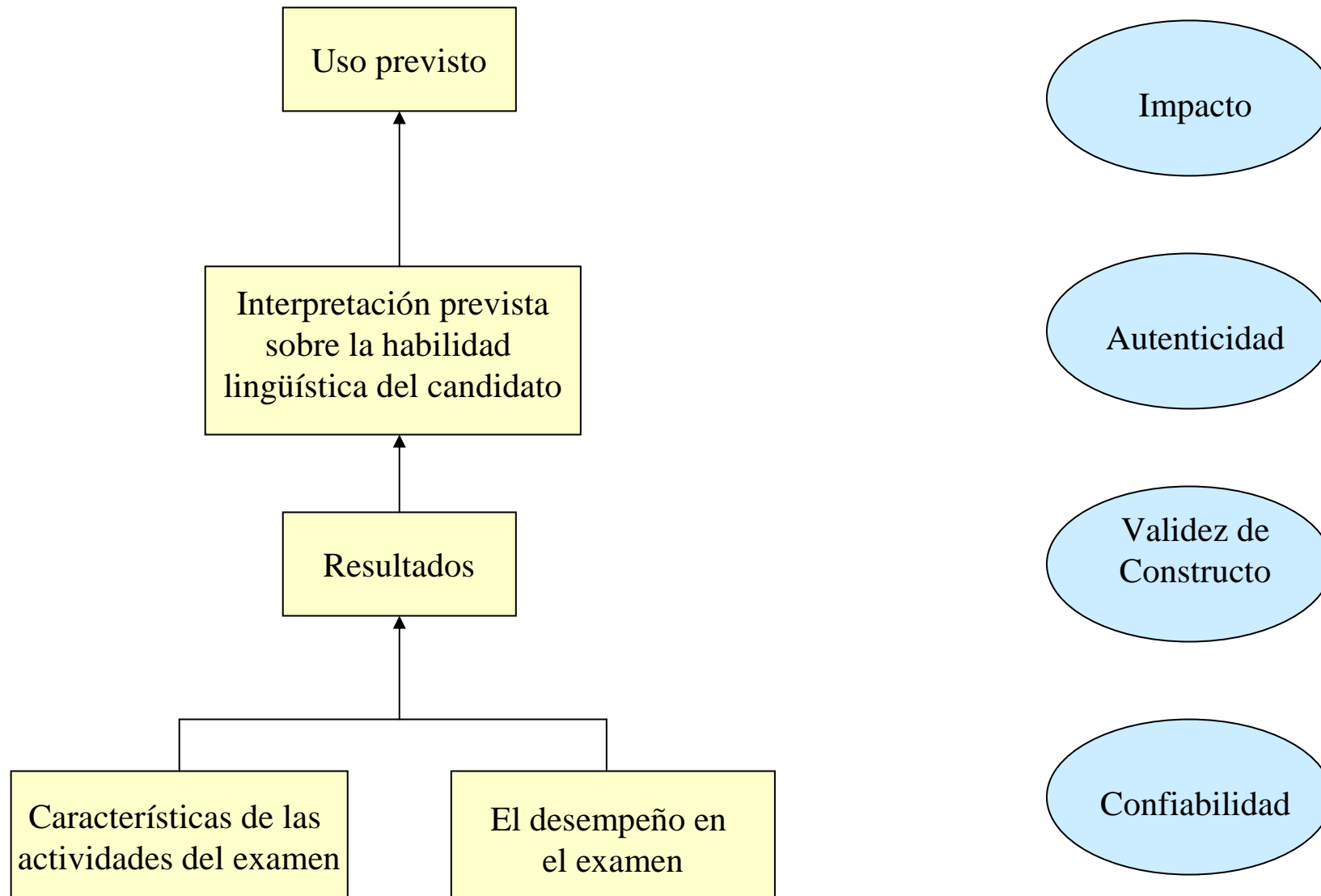
Impacto sobre los sistemas educativos y la sociedad

El uso de una evaluación de lengua tiene consecuencias para el sistema educativo y la sociedad.

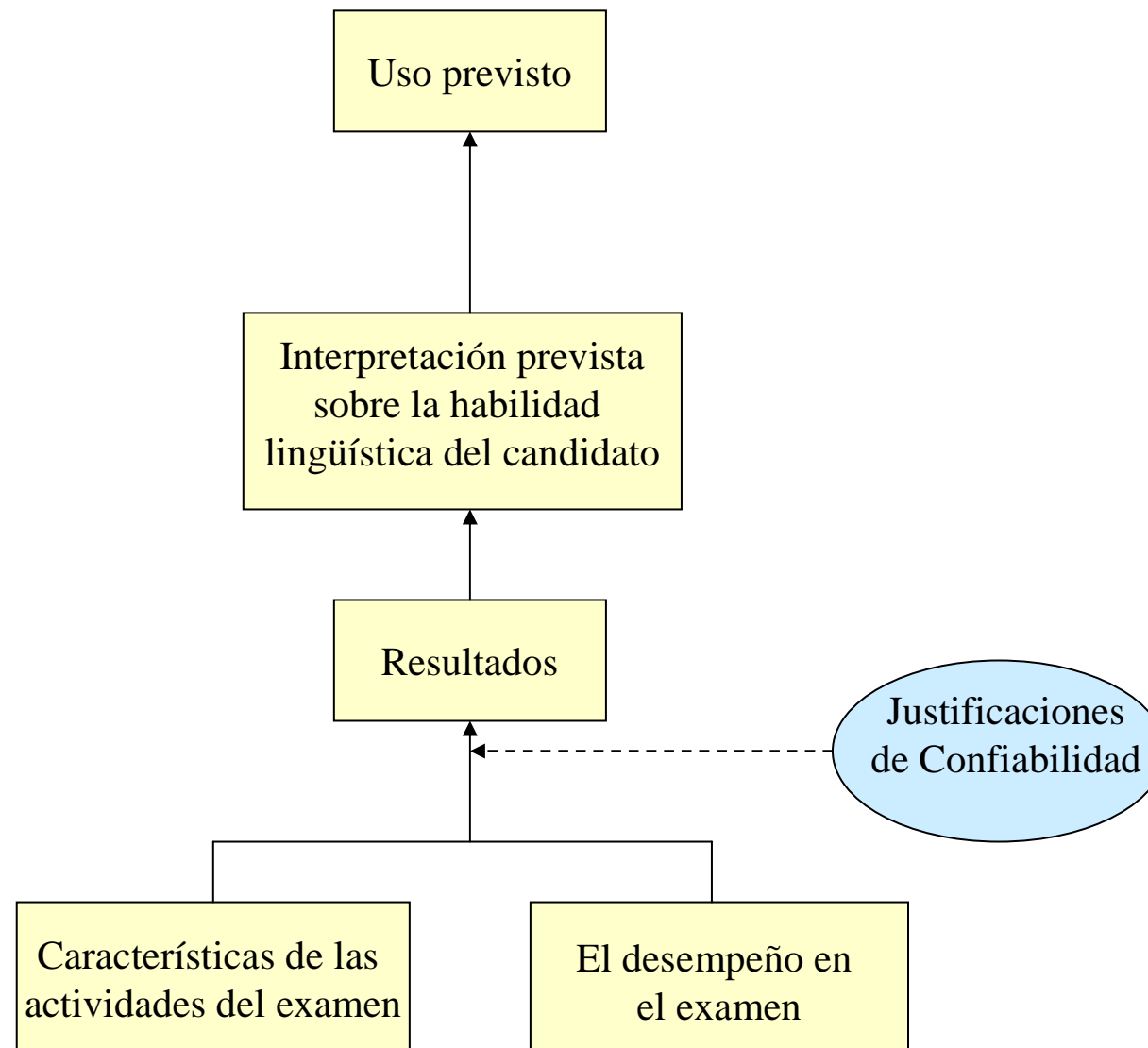
Como diseñadores y usuarios de exámenes, tenemos que considerar el sistema de valores de la sociedad y del sistema educativo que subyacen a nuestro uso del examen.

Es importante considerar que el proceso de evaluación y las decisiones tomadas sean justas.

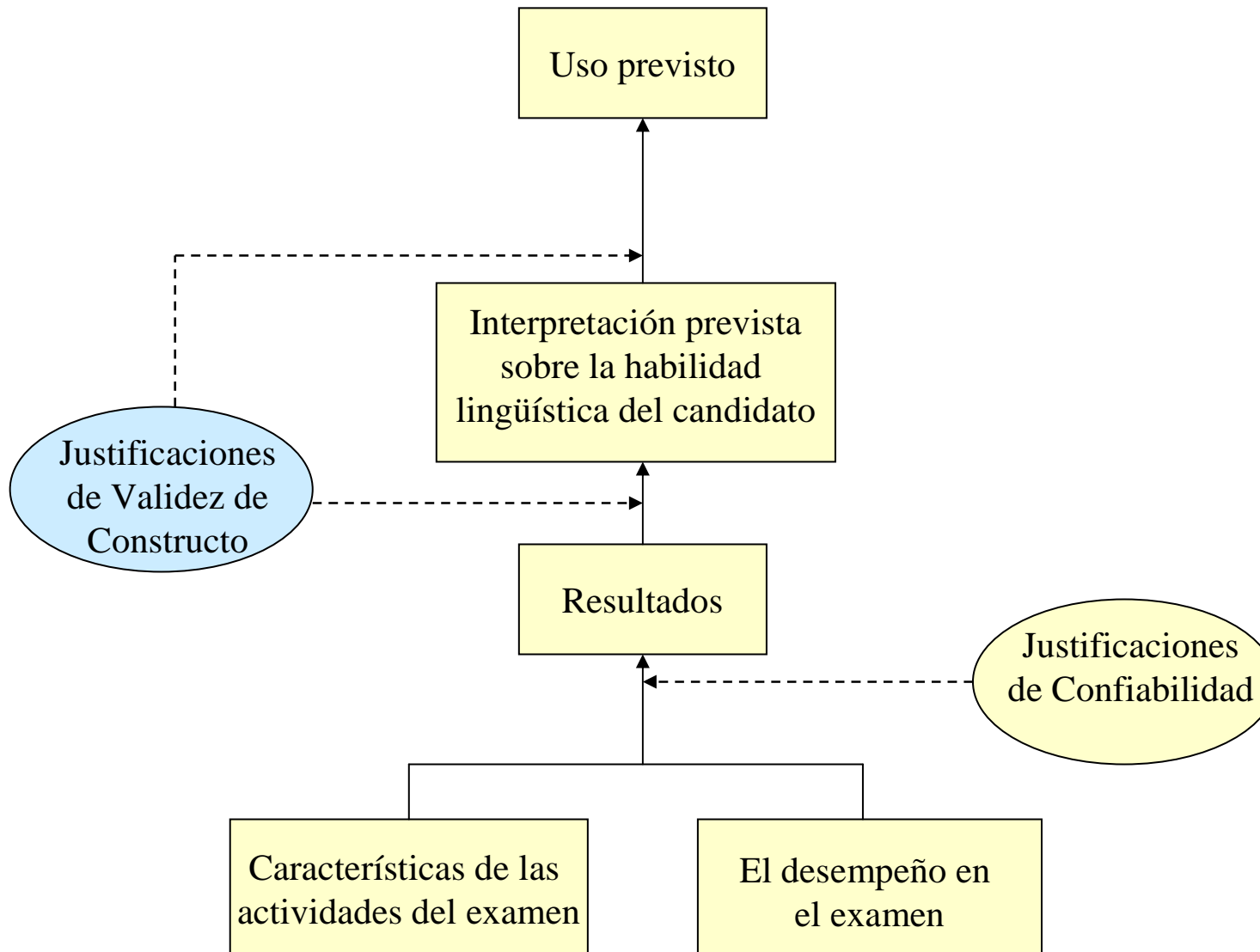
El argumento del uso de la evaluación y las cualidades de utilidad



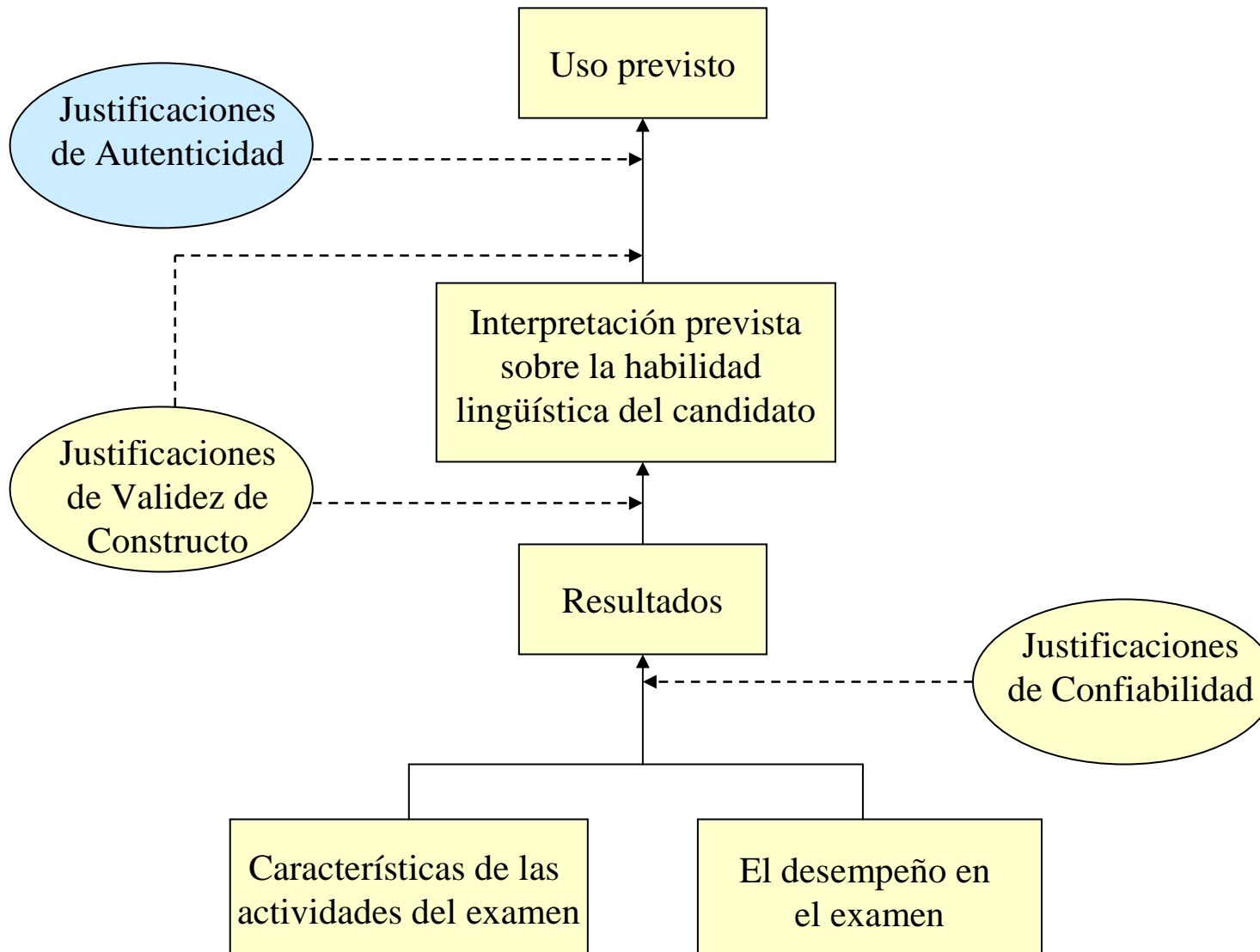
El papel de la confiabilidad en el argumento del uso de la evaluación



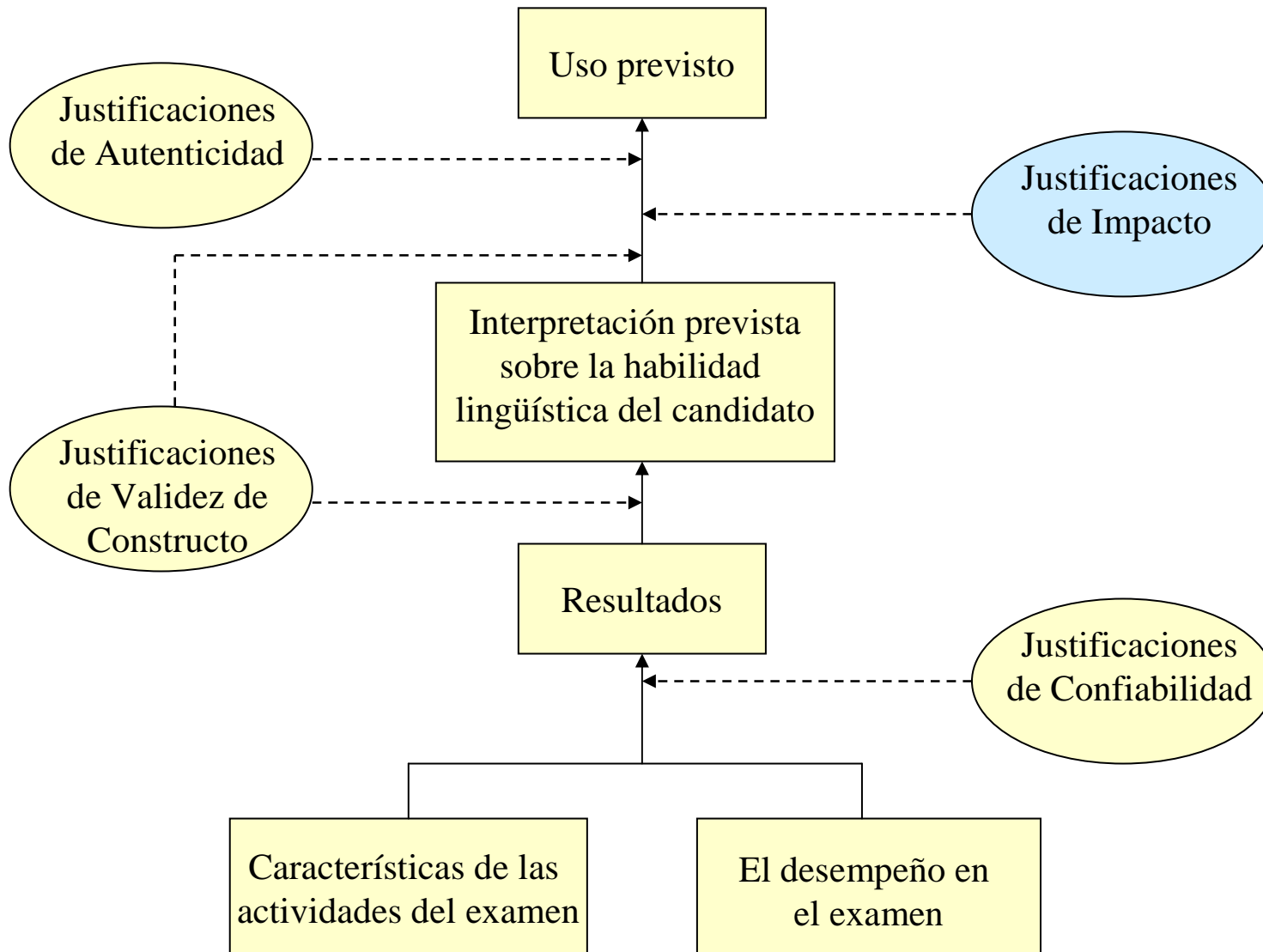
El papel de la validez de constructo en el argumento del uso de la evaluación



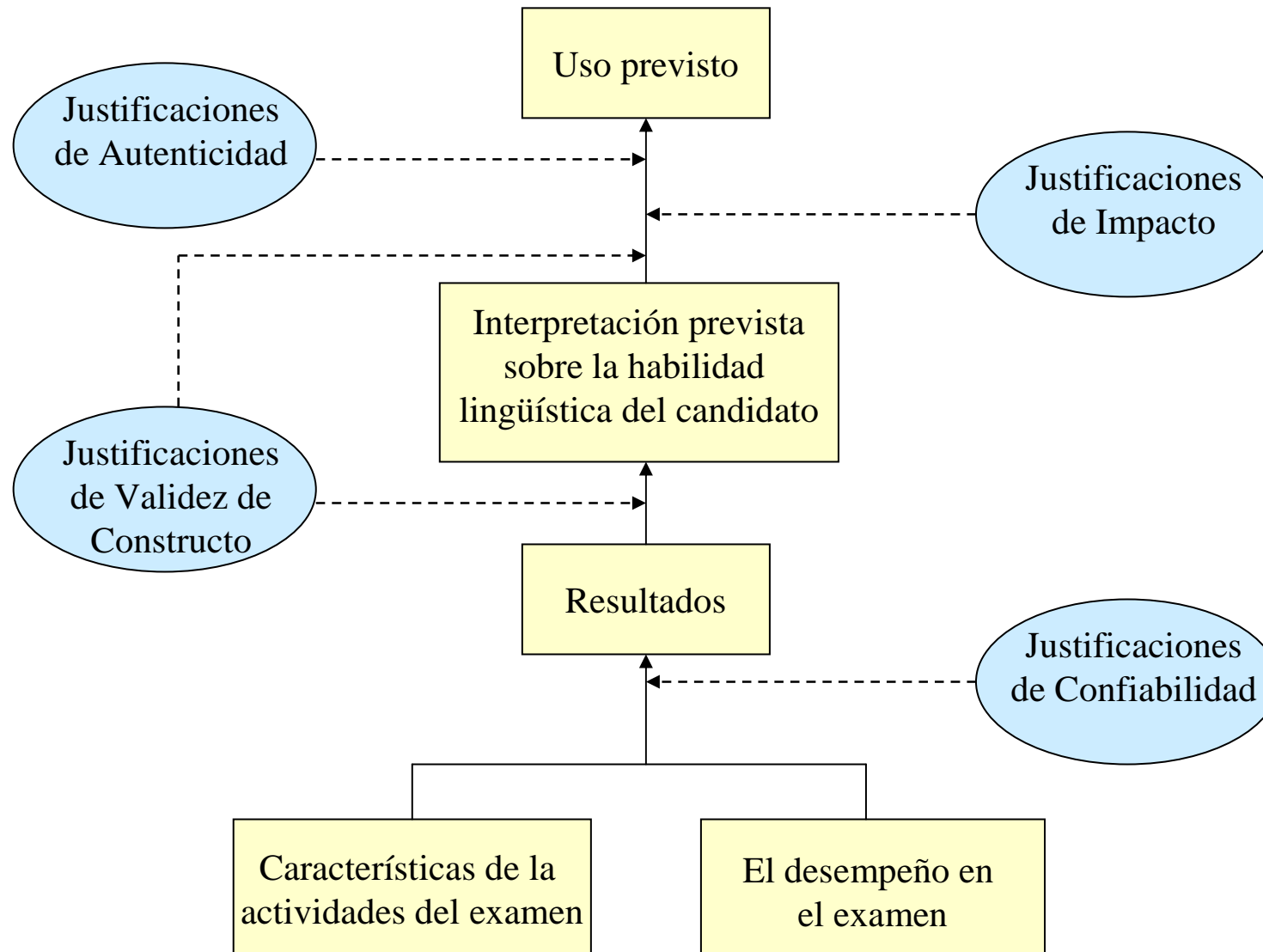
El papel de la autenticidad en el argumento del uso de la evaluación



El papel del impacto en el argumento del uso de la evaluación



El papel de las cualidades de utilidad en un argumento del uso de la evaluación



La viabilidad

- La viabilidad no es una cualidad del examen sino de todo el proceso de diseñar, desarrollar y usar un examen.
- Tiene que ver con los recursos requeridos para diseñar y desarrollar el examen y para conseguir el respaldo necesario para justificar un determinado uso del examen.
- Tenemos que focalizar nuestros recursos en las justificaciones del argumento del uso de la evaluación que requieren más apoyo y las refutaciones más problemáticas.
- Cuanto mas grande sea el impacto de la evaluación, más recursos tienen que ser invertidos para formular un argumento del uso de la evaluación y recopilar el respaldo necesario para dar apoyo a las justificaciones.

Unos puntos importantes

Es imposible demostrar que un examen es útil para el propósito previsto. Lo que sí podemos hacer es formular un argumento del uso de la evaluación claro y coherente y proporcionar evidencia de apoyo que las partes interesadas consideren convincente.

El argumento del uso de la evaluación tiene que ser revisado con regularidad para reflejar los cambios en la situación del examen. Esto es particularmente importante cuando un examen que fue diseñado originalmente para un propósito es usado para otro.

Más puntos importantes

La validación es un proceso continuo que comienza con el argumento del uso y diseño de la evaluación y continúa con la recopilación de evidencia pertinente todo el tiempo que se use el examen para tomar decisiones.

El marco teórico que nos proporciona el argumento del uso de la evaluación no es solo útil para proyectos de evaluación a gran escala. También puede ser utilizado para el diseño y desarrollo de pruebas en la clase de lengua.

El argumento del uso de la evaluación para el diseño de evaluaciones en la clase de lengua

- Usos o decisiones
- Interpretaciones sobre la habilidad lingüística del candidato
- Definición del constructo
- Identificación de las características de las actividades
- Desarrollo de un proceso de calificación
- Consideraciones de utilidad
 - Confiabilidad
 - Validez de constructo
 - Autenticidad
 - Impacto

Bibliografía

- Bachman, L. F. & Palmer, A. S. (forthcoming). *Language assessment in practice*. Oxford: Oxford University Press.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1-34.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Preguntas?

Dirección de correo electrónico: lorena.llosa@nyu.edu

Para mas información sobre el campo de evaluación de
lengua:

- ILTA (International Language Testing Association)
<http://www.iltaonline.com/>
- Language Testing Research Colloquium (LTRC)
Junio 2007, Barcelona, España