



# Evaluación como herramienta de efecto colateral (washback): ¿es esto beneficioso o dañino?

**Nick Saville**

Director, Research and Validation  
University of Cambridge  
ESOL Examinations

October 2008



# Resumen

1. Introducción
2. Efecto colateral (Washback) e impacto: Definiciones
3. El impacto y la ley de consecuencias no deseadas
4. Investigando el impacto: el enfoque de Cambridge ESOL (Inglés para hablantes de otras lenguas)
5. Conclusión



# 1. Introducción

- Exámenes
  - Acceso a oportunidades
  - Imparcialidad
- Ejercen una gran influencia sobre los procesos educativos
  - *posiblemente negativa?*
  - preocupación por las consecuencias sociales, los asuntos éticos, la confiabilidad y la transparencia



## Preguntas:

- *¿Cómo pueden los educadores y los proveedores de pruebas promover efectos beneficiosos a partir de los exámenes?*
- *¿Cómo pueden evitarse consecuencias negativas o cómo se puede mitigar su impacto?*



## ***“Impacto por Diseño” (impact by design)***

- Parte integral de un marco para el desarrollo y la validación de sistemas de evaluación
- Un concepto similar a la valoración del impacto social



## 2. Efecto colateral (Washback) e Impacto: Definiciones

- El ‘Efecto Colateral’ (*Washback* o backwash en inglés ) ha sido definido en la literatura como *el efecto de los exámenes en la enseñanza y el aprendizaje*
- Un aspecto del fenómeno más amplio conocido como *impacto* – tal como veremos más adelante



# Efecto Colateral (Washback)

- Alderson y Wall, 1993:
- 15 hipótesis acerca del efecto colateral (washback effect)
- Basadas en *quién* o *qué* puede ser afectado:
  - La enseñanza
  - El aprendizaje
  - Los contenidos
  - Los niveles de aprendizaje
  - La secuencia del proceso enseñanza/aprendizaje
  - Los grados y profundidad de la cobertura del currículo
  - Las actitudes de los docentes/estudiantes
  - Etc.





## El efecto colateral (*The washback effect*)

- Negativo?
  - Restricción de contenidos – reducción del currículo
  - Demasiado tiempo practicando para el examen
- Positivo?
  - Objetivos y resultados claros
  - Mayor motivación de los estudiantes
  - Mayor confiabilidad y transparencia de los profesores (?)
- *PERO* – pocas veces *las explicaciones de causa y efecto* son apropiadas .....



# Modelos del efecto colateral (washback models)

En la literatura de los exámenes de lengua:

- *Bailey (1996)*
- *Watanabe (2004)*
- *Cheng (2004, 2005)*
- *Green (2007)*

# Modelo de Bailey (1996) - basado en Hughes, 1993

## 3 Ps:

### Participantes

- estudiantes
- docentes

### Procesos

### Productos

- aprendizaje
- enseñanza
- materiales
- currículo

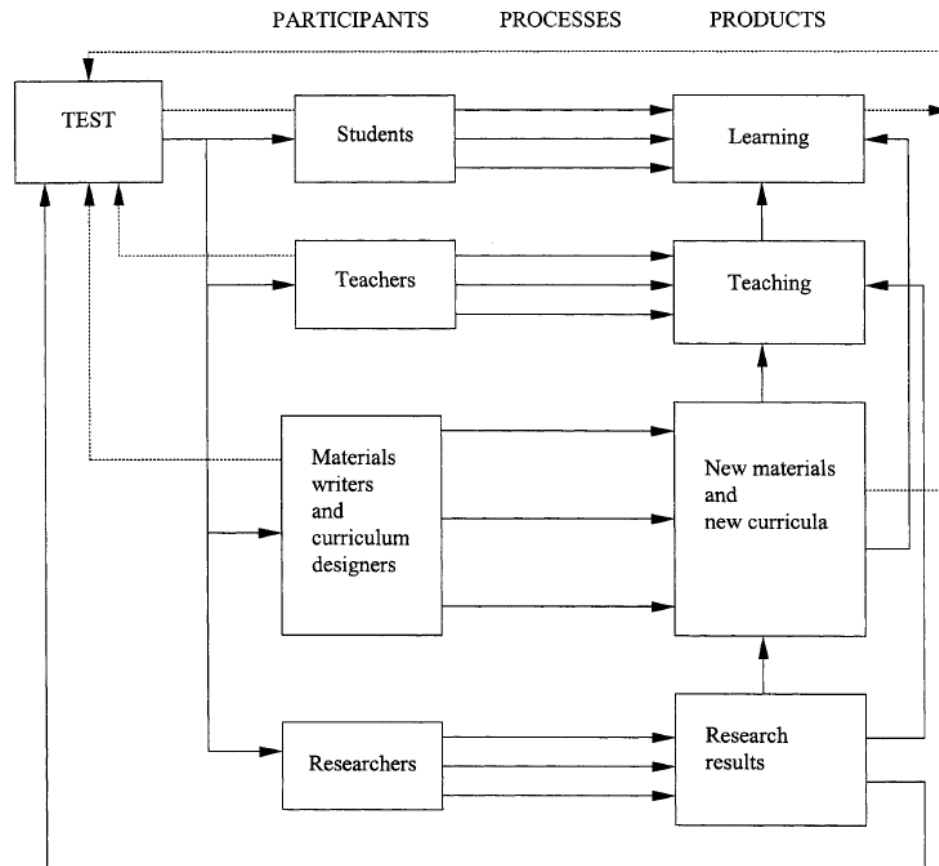
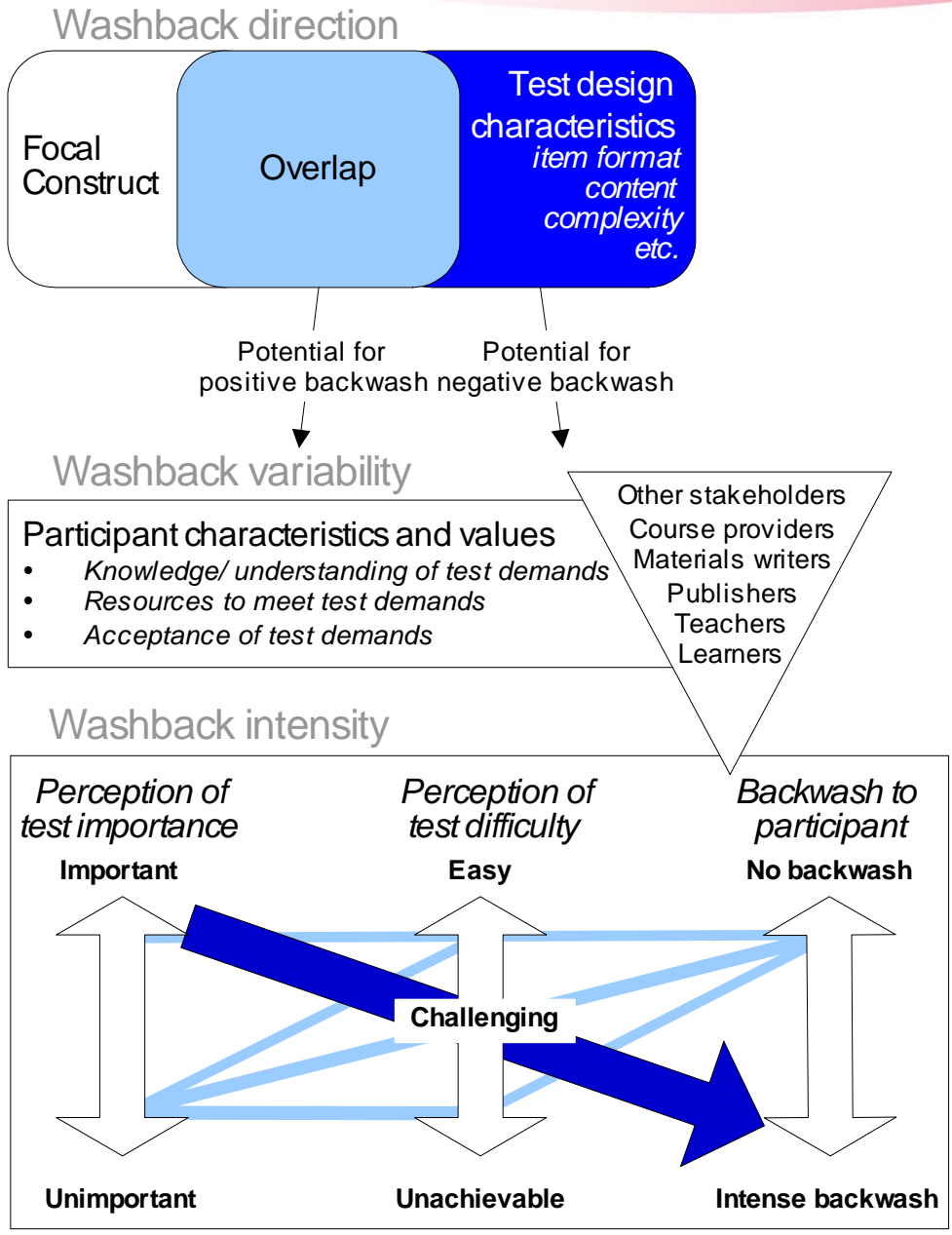


Figure 1: A basic model of washback (reprinted from Bailey, 1996a, p. 264)



# Modelo de Watanabe (2004)

- Cinco *dimensiones*:
  - *Especificidad*
  - *Intensidad*
  - *Longitud*
  - *Intencionalidad*
  - *Valor*
- Los factores relevantes al *proceso* de la *influencia* de los exámenes se relacionan con:
  - *La prueba o test*
  - *El estatus*
  - *Las partes interesadas*



# Modelo de Green

Ver: Studies in Language Testing, 25, 2007: IELTS Washback in Context

Efecto Colateral (Washback) será **más intenso** – tendrá los efectos más poderosos sobre los comportamientos de enseñanza y aprendizaje – cuando los participantes vean **la prueba como un reto** y los **resultados como algo importante** (altos intereses)



# El impacto

El impacto está relacionado con los efectos y las consecuencias que una prueba puede tener *fuera del salón de clase* y con el contexto de aprendizaje inmediato:

- En las oportunidades de desarrollo profesional y de vida
- En los sistemas educativos y en la sociedad



# El Impacto

- El impacto maneja influencias mayores e incluye los “*macro contextos*” – las pruebas y exámenes en la sociedad
- El efecto colateral (washback) es un aspecto del impacto relacionado con los “*micro contextos*” del aula y el colegio



# El Impacto

- Investigar el impacto es una parte importante para la validación
  - un componente esencial para establecer la **utilidad de un sistema de evaluación** – aptitud para fines específicos y contextos de uso
- Consistencia con los puntos de validez de Messick (1989, 1996)
  - *Consecuencias de la validez*
  - Imparcialidad y ética

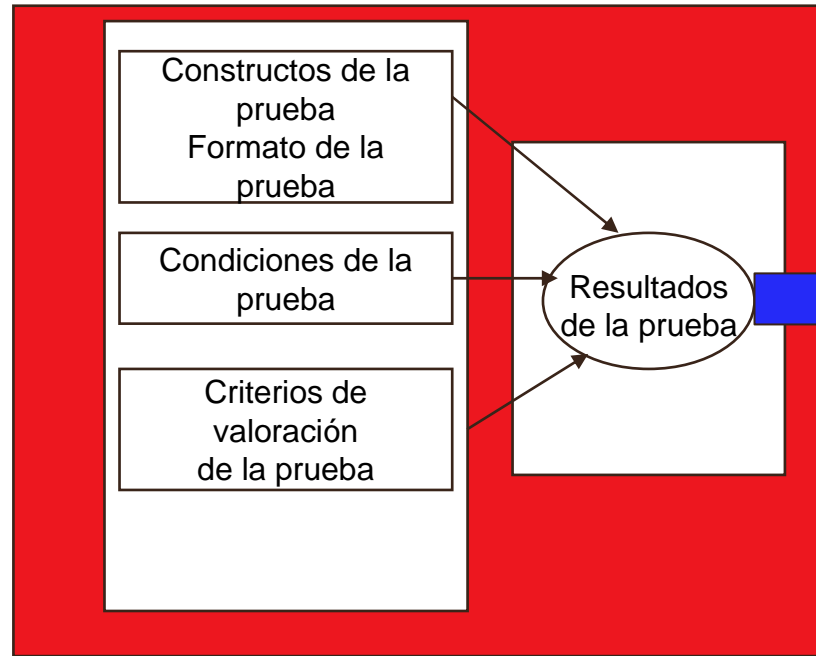


# Impacto

- Sistemas Educativos – fenómenos complejos
- Circunscripciones de las partes interesadas
  - Diferentes partes interesadas en los procesos educativos
  - Red de relaciones compleja

# Las partes interesadas en la circunscripción de pruebas

Estudiantes  
Docentes  
Constructores de pruebas  
y Examinadores  
Instituciones receptoras  
Propietarios de colegios  
Futuros empleadores  
Agencias  
gubernamentales  
Agentes profesionales  
Centros Administradores  
de pruebas  
Diseñadores de materiales  
Editoriales  
etc



Sistema de Evaluación

Estudiantes  
Padres/Acudientes  
Docentes  
Instituciones  
receptoras  
Empleadores  
Propietarios de  
colegios  
Examinadores  
Agencias  
gubernamentales  
Agentes  
profesionales  
Investigadores  
académicos  
Constructores de  
pruebas y  
Examinadores  
etc

**Aportes al diseño de pruebas**

**Contextos del uso de pruebas - consecuencias**



# Impacto

- Voces múltiples
  - Puntos de vista, creencias y actitudes
- Es importante para los proveedores de pruebas
  - Obtener retroalimentación de las partes interesadas
  - Tener en cuenta los puntos de vista de las partes interesadas
- Relación dinámica entre los micro y los macro contextos



### 3. La ley de consecuencias no deseadas

- *“Cualquier acción intencionada producirá consecuencias no deseadas” o efectos secundarios*
- **“La ley de Goodhart”** (o la ley de Campbell en USA)
  - Una variante de la ley de consecuencias no deseadas



## “La ley de Goodhart”

- *“Todos los indicadores de desempeño pierden su significado cuando se adoptan como objetivos de la política”*

Ejemplos:

- Inglaterra – objetivos de rendimiento escolar – *tablas de clasificación de colegios y escuelas (school league tables)*
- USA – Qué Ningún Niño se Quede Atrás
  - No Child Left Behind Act = NCLB por sus iniciales en inglés
- *“Entre más claridad tengas en cuanto a lo que quieres, mayor posibilidad tienes de conseguirlo - pero menos posibilidades tendrás de que signifique lo que inicialmente querías”*

(Dylan Wiliam, 2008)



## Incentivos perversos? (*perverse incentives*)

- La política de evaluación puede crear una tensión entre:
  - Los objetivos educativos en el *micro contexto* (la enseñanza y el aprendizaje en las escuelas) y
  - Un requisito de rendición de cuentas en el *macro contexto*



## ¿Qué papel pueden jugar los proveedores de pruebas?

¿Cómo pueden los proveedores de pruebas asegurarse de que:

- *los sistemas de evaluación son "aptos para el propósito"?*
- *la investigación se realiza con el objetivo de determinar lo que ocurre en contextos reales de uso?*
- *se realizan acciones de mitigación de efectos negativos cuando estos ocurran?*



## 4. Investigando el impacto

- Cambridge ESOL
  - Una agencia internacional de evaluación
- **Máximas del Impacto de la evaluación**
  - Un intento de integrar un enfoque orientado a la acción para investigar el impacto en las prácticas de trabajo
  - Milanovic and Saville, 1996

# Máximas para alcanzar / monitorear el impacto

## Máxima 1 PLANEE

Use un enfoque lógico y explícito para el *desarrollo de pruebas*

## Máxima 2 APOYE

Apoye a las *partes interesadas* en el proceso de evaluación

## Máxima 3 COMUNIQUESE

Proporcione *información transparente*, amplia y útil

## Máxima 4 MONITOREE y EVALUE

Recopile todos los *datos relevantes* y analícelos según sea necesario.

Milanovic and Saville, 1996 - *Considering the impact of the Cambridge examinations*



# Máxima 1 Planee

Use un enfoque lógico y explícito para el desarrollo de pruebas

- Un *modelo* cíclico y reiterativo
- crea las *condiciones necesarias* para el desarrollo apropiado de pruebas y para que haya validación continua
- Inicia con el propósito de evaluación – incluyendo la previsión
  - la forma en que la prueba debe (o puede) ser utilizada
  - la pertinencia y utilidad – consecuencias sociales e implicaciones de valor
  - Posibles efectos colaterales (no planeados)



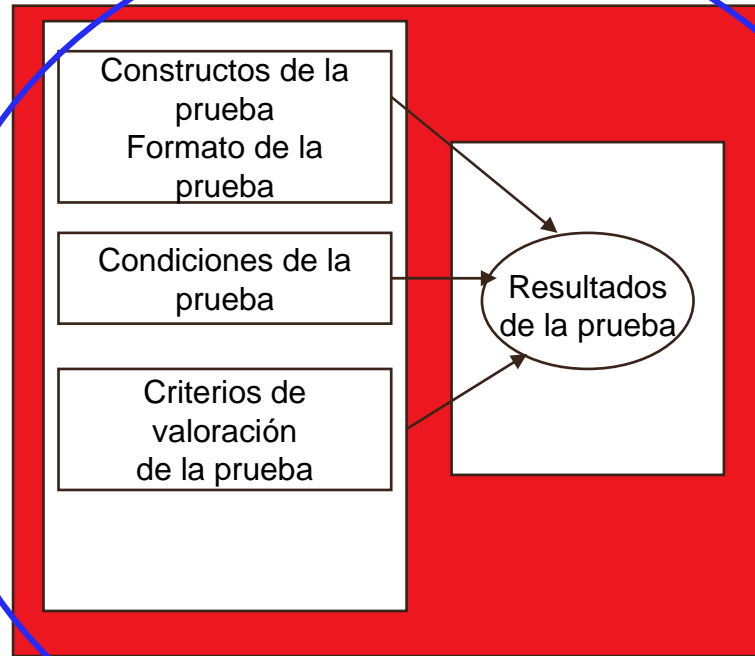
## Maxima 2 Apoye

### Apoye a las partes interesadas en el proceso de evaluación

- Participación de los interesados - durante el diseño y desarrollo de la prueba
- Programas de apoyo profesional
- Formación de personal idóneo para trabajar en todos los aspectos del ciclo de evaluación

# Las partes interesadas en la circunscripción de pruebas

Estudiantes  
Docentes  
Constructores de pruebas y Examinadores  
Instituciones receptoras  
Propietarios de colegios  
Futuros empleadores  
Agencias gubernamentales  
Agentes profesionales  
Centros  
Administradores de pruebas  
Diseñadores de materiales Editoriales  
etc



Sistema de Evaluación

Estudiantes  
Padres/Acudientes  
Docentes  
Instituciones receptoras  
Empleadores  
Propietarios de colegios  
Examinadores  
Agencias gubernamentales  
Agentes profesionales  
Investigadores académicos  
Constructores de pruebas y Examinadores  
etc

**Aportes al diseño de pruebas**

**Contextos del uso de pruebas - consecuencias**



## Máxima 3 Comuníquese

Proporcione *información* transparente, amplia y útil

- Explicar aspectos relativos a la evaluación adaptados para las diferentes partes interesadas
- Un gran reto para todos los proveedores de evaluaciones!



## Máxima 4      Monitoree y Evalúe

Recopile todos los datos relevantes y analícelos según sea necesario.

- Por ejemplo, la retroalimentación de los candidatos y sus profesores
  - de los contextos locales, donde la prueba se utiliza (es decir a nivel micro)
- Empeño a largo plazo, ya que incluye:
  - el desarrollo de **instrumentos adecuados** para la recolección de datos adecuados
  - **metodologías de investigación** pertinentes
- Evalúe rutinariamente la utilidad de la prueba
  - determine la necesidad de cambios y revisiones periódicas



## Resultados de estudios de impacto

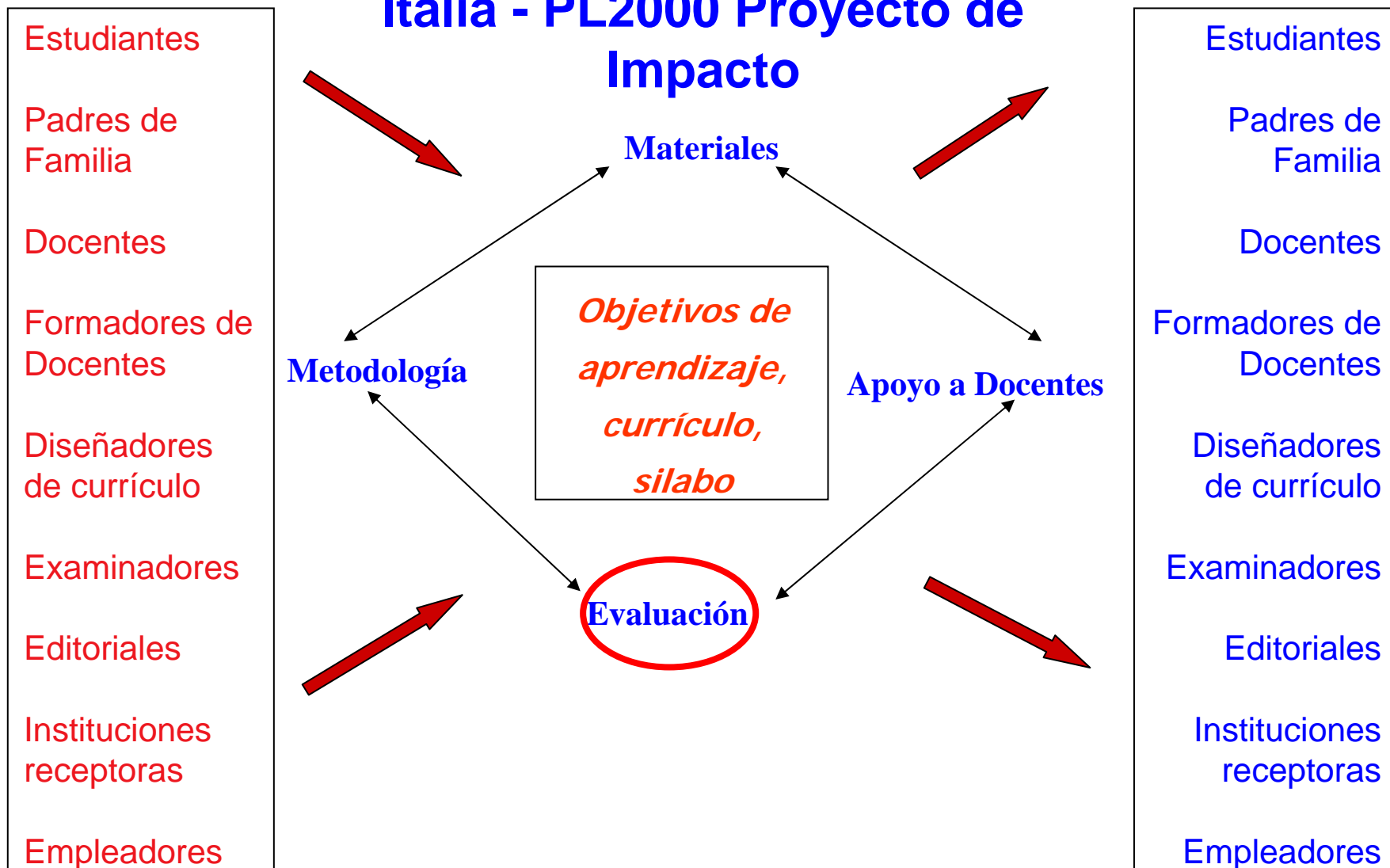
- La encuesta mundial del **impacto de IELTS**
- El *Progetto Lingue 2000* italiano (proyecto PL2000)
- Ver Hawkey, (2006)
  - Studies in Language Testing, 24  
*The theory and practice of impact studies: Messages from studies of the IELTS test and Progetto Lingue 2000*



# Progetto Lingue 2000

- La política era
  - “... para introducir la innovación en la enseñanza y el aprendizaje de otros idiomas, poniendo un mayor énfasis en el desarrollo de la competencia comunicativa en todos los grados del sistema escolar ”
- Este incluyó:
  - La adopción del Marco Común Europeo de Referencia (CEFR) como base para *objetivos* y estándares de aprendizaje
  - certificación *de proficiencia* – la evaluación (por un ente certificador internacionalmente reconocido)

# Italia - PL2000 Proyecto de Impacto





## 5. Conclusiones

- La Investigación de Impacto ahora refleja la creciente importancia de enfoques *basados en evidencia* para la planificación y la evaluación educativa
- Adoptando un enfoque *“impacto por diseño”* por medio de estrategias de desarrollo y validación de pruebas, y utilizando los hallazgos de la investigación de impacto para guiar futuras acciones – políticas y prácticas de evaluación mas eficientes se pueden desarrollar para alcanzar las necesidades de la educación contemporánea



# Gracias



# Evaluación como herramienta de efecto colateral (washback): ¿es esto beneficioso o dañino?

**Nick Saville**

Director, Research and Validation  
University of Cambridge  
ESOL Examinations

October 2008