

Chi ha paura di Esther Duflo?

La sfida della valutazione sperimentale



Inclusa dall'*Economist* tra gli 8 economisti più influenti al mondo, francese, ha meno di 40 anni e ha avuto un'influenza determinante nella diffusione del metodo sperimentale allo scopo di testare l'efficacia degli aiuti allo sviluppo

Perché e quando usare il metodo sperimentale?

Una ragione preliminare la si trova in questa vignetta



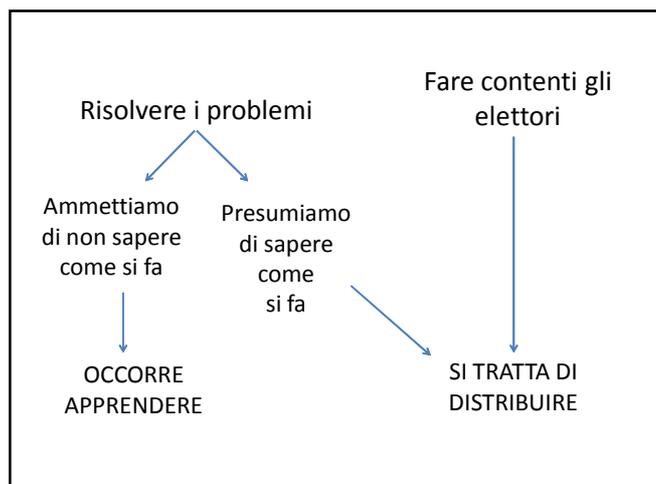
— Sia più preciso, signore: il regalo deve rendere felice la persona che lo riceve, oppure deve risolvere un problema?

Uno dei grandi dilemmi delle democrazie moderne:

risolvere i problemi
o
fare contenti gli elettori?

Nel caso raro ci sia una reale volontà di risolvere i problemi, il dilemma si sposta su un altro piano

imparare a risolvere i problemi
o
presumere di sapere come si fa



Cosa rende meno risibile del solito questo ragionamento?

Non resta più molto da distribuire

Forse è tempo di capire se sono soldi ben spesi

Prendiamo uno degli innumerevoli casi in cui vi è un alto rischio di sbagliare a causa dell'insufficiente conoscenza sui legami di causa-effetto

Poniamo per ipotesi...

Che il Ministero trasferisca alle Regioni tutte le competenze in materia di istruzione elementare
Tocca ora al Consiglio Regionale legiferare su una materia delicata: la dimensione delle classi nelle scuole elementari

...come si stabilisce qual'è la dimensione ideale di una classe di scuola elementare?

Si mescolano considerazioni diverse (costo, equità, efficacia)

I risvolti importanti della decisione sono:

La dimensione media delle classi influisce su quanto il sistema scolastico può permettersi di pagare gli insegnanti

Non è un processo automatico, ma è ovvio che un modo per pagare di più gli insegnanti è chiedere loro di lavorare con più alunni

L'obiezione principale è che classi più grandi riducono le chance di apprendimento da parte degli studenti

Lo diceva anche mia nonna, ma siamo sicuri sia vero?

Evidence-based policy

Ci vuole evidenza del fatto che l'apprendimento peggiori con la dimensione della classe in modo da giustificare la preferenza per classi piccole.

Potrebbe valere per certi intervalli e non per altri, potrebbe valere per alcuni tipi di studenti e non per altri

Un caso concreto

Confrontiamo le classi di 17 allievi di 5° elementare con le classi di 23 allievi.
Cosa ne ricaviamo?

Quattro regioni commissionano altrettanti studi, che utilizzano campioni diversi (ma tutti di 1000 alunni, metà in classi di 17 e metà in classi di 23)

Come misura di apprendimento si usa un test Invalsi, perfetto, onnicomprensivo e amato da tutti. Il punteggio al test va da 0 a 100

I quattro studi ottengono i seguenti risultati

studio	Punteggio Invalsi		Differenza
	classi da 17 alunni	classi da 23 alunni	
I	48	60	-12
II	52	52	0
III	56	44	12
IV	60	36	24

Come è possibile tanta variabilità?

Una spiegazione intuitiva è la seguente

Mettiamo esistano due tipi di alunni

Alunni B e alunni A

La verità (che non possiamo osservare) è che Gli alunni B hanno lo stesso rendimento che siano in una classe da 17 o in una da 23

Gli alunni A beneficiano invece molto dall'essere in una classe più piccola, raddoppiando il punteggio Invalsi

	Classi piccole	Classi grandi	Effetto	Percentuale
Alunni B	60	60	0	70%
Alunni A	40	20	20	30%

Il risultato del confronto classi grandi e piccole dipende largamente da come i B e gli A sono distribuiti tra i due tipi di classi

Gli asini nei pollai!

	Classi piccole	Classi grandi	totale
Alunni B	500	200	700
Alunni A	0	300	300
totale	500	500	1000
Punteggio Invalsi	60	36	24

Nessun asino nei pollai

	Classi piccole	Classi grandi	totale
Alunni B	200	500	700
Alunni A	300	0	300
totale	500	500	1000
Punteggio Invalsi	48	60	-12

La dimensione forse non fa differenza

Casi meno estremi

	Classi piccole	Classi grandi	totale
Alunni B	400	300	700
Alunni A	100	200	300
totale	500	500	1000
Punteggio Invalsi	52	52	zero

O forse si

	Classi piccole	Classi grandi	totale
Alunni B	300	400	700
Alunni A	200	100	300
totale	500	500	1000
Punteggio Invalsi	56	44	+ 12

Il metodo sperimentale serve ad evitare questa instabilità di risultati

Il passo essenziale è assegnare mediante sorteggio gli alunni alle classi, in modo da evitare ogni selezione sistematica

1000 alunni vengono assegnati a caso, 500 alle 22 classi grandi e 500 alle 28 classi piccole

Randomizzando si ottengono due gruppi equivalenti

	Classi piccole	Classi grandi	totale
Alunni B	350	350	700
Alunni A	150	150	300
totale	500	500	1000
Punteggio Invalsi	54	48	+ 6

La stima di 6 punti è formalmente corretta, ma non è interessante da un punto di vista di policy, in quanto si riferisce ad un effetto medio che nasconde una notevole eterogeneità

Anche se l'esperimento da solo non riesce a cogliere questo fenomeno, evita però le conclusioni più strampalate: nell'esempio la stima dell'effetto medio è a metà tra le due stime intermedie (zero e 12 punti)

L'idea è quella di eliminare tutte le differenze tra chi riceve e chi non riceve una prestazione, in modo da isolare l'effetto che cerchiamo

La randomizzazione non risolve tutti i problemi ma evita di fare errori clamorosi e dare cattivi consigli

Soprattutto evita di chiamare effetto le differenze dovute a qualcosa d'altro

Il caso forse più famoso di esperimento nella scuola

Lo STAR experiment

(Students Teachers Achievement Ratio)

realizzato tra il 1985 e il 1989 nelle scuole elementari del Tennessee, coinvolgendo 11600 alunni K-3 e 1300 insegnanti di 79 scuole.

La coorte nata nel 1980 inizia il Kindergarten (obbligatorio) nel 1985

gli alunni nelle scuole che partecipano vengono assegnati mediante sorteggio a tre regimi:

- Classi 13-17 alunni
- Classi 23-26 alunni
- Classi 23-26 alunni con insegnante d'appoggio

Restano in queste classi fino alla fine della 3° elementare. Ogni anno si misura l'apprendimento mediante test standardizzati, ma poi si misura il tasso di drop-out, il passaggio al college, il reddito, i problemi con la giustizia..

I risultati sono molto chiari:

- L'insegnante d'appoggio non fa nessuna differenza
- le classi 13-17 hanno punteggi di lettura e matematica moderatamente migliori delle classi di 23-26 alunni
- Le differenze forti si hanno per tra gli studenti afro-americani (che guadagnano molto più dei bianchi dall'essere in una classe piccola)
- Una differenziale minore si ha confrontando studenti poveri e non (free lunch program)
- Il guadagno tende ad essere permanente nel tempo.
- Un'analisi costi benefici mostra un tasso di rendimento del 5% all'anno dall'investimento necessario a ridurre le classi, assumendo altri insegnanti.

Ancora un fatto significativo

L'iniziativa è partita ed è stata finanziata dalla State Legislature (CONSIGLIO REGIONALE!) del Tennessee.
(27 ANNI FA!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!)

L'esperimento è costato 12 milioni di dollari, circa 9 milioni di euro

Sui risultati dell'esperimento STAR sono stati condotte centinaia di analisi, che continuano tuttora, seguendo gli alunni ora > 30enni

L'iniziativa è partita ed è stata finanziata dalla State Legislature (CONSIGLIO REGIONALE!) del Tennessee

27 ANNI FA!