

# Esperimenti paradigmatici: il gioco dell'Ultimatum<sup>+</sup>

Francesco Guala<sup>\*</sup>

[francesco.guala@unimi.it](mailto:francesco.guala@unimi.it)

## ABSTRACT

The Ultimatum Game is one of the most successful experimental designs in the history of the social sciences. In this paper I try to explain this success what makes it a <sup>3</sup>paradigmatic experiment<sup>2</sup> stressing in particular its versatility. Despite the intentions of its inventors, the Ultimatum Game was never a good design to test economic theory, and is now mostly used as a heuristic tool for the observation of non-standard preferences or as a <sup>3</sup>social thermometer<sup>2</sup> for the observation of culture-specific norms.

---

## 1. INTRODUZIONE

Immagina di avere appena ricevuto 10 euro. Questa somma dovrà essere divisa con un partner anonimo e invisibile, e dovrai accordarti con lui (o lei) su come dividerla. Lo spazio per la discussione è quasi inesistente: potrai soltanto offrire una divisione della torta, e il tuo partner potrà solo accettare o rifiutare. Se la rifiuta, perderete entrambi l'opportunità di dividervi i 10 euro; se l'accetta, vi porterete a casa la vostra porzione, come indicato dalla divisione che hai proposto.

Questa è sostanzialmente la situazione strategica nota come "gioco dell'Ultimatum" (GU). Probabilmente avrai già incontrato il GU. Se, come Robert Axelrod ha notato, il Dilemma del Prigioniero è diventato l'*E. coli* delle scienze sociali, allora il GU ne è senz'altro la *Drosophila melanogaster*. La battuta di Axelrod va presa sul serio. Suggerisce che la teoria dei giochi sperimentale ha raggiunto un livello di maturità paragonabile a quello di scienze sperimentali affermate quali la biologia molecolare. Suggerisce anche che questa maturità si manifesta in particolare nell'emergere e nel consolidamento di prototipi sperimentali o "esperimenti paradigmatici" analoghi agli "organismi modello" della biologia sperimentale. Gli esperimenti paradigmatici hanno varie caratteristiche epistemiche distintive. In primo luogo, anni di sperimentazione fanno sì che li conosciamo molto bene, sicuramente meglio di qualsiasi altro sistema nel dominio di applicazione della scienza. In secondo luogo, essi vengono usati per formulare una varietà di inferenze ad altri sistemi che sono meno accessibili allo studio sperimentale. E infine, sono solitamente strumenti versatili che possono essere utilizzati per diversi scopi in diversi momenti e in diversi contesti di investigazione scientifica. In questo articolo mi concentrerò principalmente su questa caratteristica, ma anche le altre avranno un ruolo nella storia del GU.

---

<sup>+</sup> Una versione in inglese di questo articolo è apparsa sulla rivista *Philosophy of Science* 75 (2008) pp. 658-669. Marcel Boumans, Margaret Schabas, Bob Sugden, e diversi partecipanti alla conferenza della Philosophy of Science Association a Vancouver nel Novembre 2006 hanno fornito diversi commenti e suggerimenti. Tutti gli errori rimanenti sono a mio carico.

<sup>\*</sup> Dipartimento di Economia Università degli Studi di Milano



## 2. CONTROLLARE LE TEORIE

Il GU fu concepito e utilizzato per la prima volta da un gruppo di economisti sperimentali tedeschi guidati da Werner Güth nei primi anni '80 del Novecento (Güth, Schmittberger and Schwarze 1982). Güth e colleghi intendevano studiare la contrattazione sequenziale, e scelsero il GU principalmente perché è il problema di contrattazione sequenziale più semplice che si possa immaginare.<sup>1</sup> La semplicità era volta a minimizzare i costi cognitivi di calcolo. È risaputo che gli individui faticano ad analizzare i giochi dinamici complessi, in particolare quando devono operare la cosiddetta “induzione all’indietro” (*backward induction*). Nel GU, non si può non capire che il gioco finisce dopo la mossa del secondo giocatore, per cui il rumore nei dati dovuto alla mancata comprensione della decisione sperimentale dovrebbe essere ridotto al minimo.

Il team di Güth intendeva anche controllare alcuni risultati di esperimenti precedenti. Fouraker e Siegel, due psicologi, avevano realizzato negli anni '50 una serie di studi della contrattazione ormai divenuti classici. Güth e i suoi colleghi trovavano sorprendente che le offerte inique venissero rifiutate raramente negli esperimenti di Fouraker e Siegel. Ipotizzarono che questo potesse essere dovuto ad alcuni dettagli del loro disegno, e decisero di controllarli modificando leggermente la situazione sperimentale.

In effetti, le offerte inique sono rifiutate nel GU. Nel loro primo studio, Güth, Schmittberger e Schwartz scoprirono che in media il primo giocatore offre il 35% della torta. Osservarono una risposta modale al 50% e pochi rifiuti da parte del secondo giocatore (circa il 10%). Quando il gioco venne ripetuto una settimana più tardi (condizione con “giocatori esperti”) l’offerta media scese al 31%. L’offerta modale al 50% scomparve, con la maggior parte delle offerte ora fra il 20 e il 30%. C’erano anche più rifiuti con la ripetizione (circa il 30%). Da allora, questi risultati sono stati replicati numerosissime volte, e costituiscono la cosiddetta “anomalia del GU”. Ma anomalia rispetto a che cosa?

## 3. CONTROLLARE LA TEORIA E MANIPOLARE LE PREFERENZE

Secondo la teoria dei giochi classica, il primo giocatore dovrebbe offrire quasi nulla e il secondo dovrebbe accettare. L’idea è che il secondo giocatore dovrebbe affrontare una decisione apparentemente banale: non portarsi a casa niente oppure ciò che è stato offerto dal primo giocatore. Supponiamo che il minimo che si può offrire sia 1 euro. Un euro è meglio di niente, quindi il secondo giocatore dovrebbe accettare. Assumendo la conoscenza comune del gioco e della razionalità, la divisione 1/9 è l’unico equilibrio di Nash del GU.

Questa “previsione standard” per la verità è derivata per mezzo di un meccanismo abbastanza complesso: da una teoria del gioco strategico (il “nucleo” della teoria dei giochi) più un insieme di assunzioni riguardo alle preferenze e alle credenze dei giocatori. Secondo l’assunzione dell’egoismo, i giocatori preferiscono avere più denaro che meno, e non badano ai guadagni altrui. Il GU dunque è un’anomalia rispetto a una previsione derivata da un modello composito (il modello della “razionalità egoista”) e solleva il noto problema di Duhem-Quine. La teoria del gioco strategico (l’assunzione di razionalità) è responsabile dell’anomalia, l’assunzione dell’egoismo è inadeguata a catturare il comportamento nel GU, le credenze

---

<sup>1</sup> La contrattazione è solitamente rappresentata nella teoria dei giochi come un problema di suddivisione, dove il surplus generato dallo scambio fra due beni deve essere allocato fra le parti. I 10 euro nel GU rappresentano in qualche modo questo surplus.



individuali divergono dalle assunzioni classiche, oppure qualche altro elemento del modello non è in grado di spiegare ciò che succede in un GU sperimentale?

La teoria della razionalità che costituisce il nucleo della teoria dei giochi è una teoria del tipo “se... allora...”: dice che se le loro preferenze e credenze sono così-e-così, allora gli individui si comporteranno così-e-così. Se le preferenze non sono egoistiche come postulato dal modello classico, quindi la teoria della razionalità non può essere controllata in un gioco *singolo* come il GU. La teoria dei giochi sperimentale “kosher” dovrebbe dunque cominciare con una misurazione delle preferenze in circostanze non-strategiche, e proseguire controllando se la teoria del gioco strategico formula previsioni corrette utilizzando le misurazioni preliminari come dati iniziali nella procedura di controllo.

Tuttavia la maggior parte della teoria dei giochi sperimentale non è “kosher” da questo punto di vista. La maggior parte degli esperimenti comincia *postulando* il contenuto delle preferenze individuali, invece di individuarle empiricamente. Il GU, come abbiamo visto, non fa eccezione. Ma se il GU non controlla davvero la teoria del gioco strategico (il “nucleo” della teoria dei giochi) perché è diventato così famoso e ampiamente replicato?

Come il Dilemma del Prigioniero, il GU è così semplice che nessuno dubita davvero che i soggetti sperimentali siano razionali, nel senso minimale che le loro azioni seguono dalle loro preferenze e credenze. Piuttosto, il GU è interessante perché rende osservabili le *preferenze non-standard*. Intuitivamente sembra plausibile che il secondo giocatore possa preferire di rinunciare a un po' di soldi per punire le offerte inique, e che il primo preveda che un'offerta iniqua possa urtare i sentimenti del secondo giocatore. I secondi giocatori hanno preferenze “rivolte agli altri”, influenzate forse da una preoccupazione per l'equità. Infatti, se consideriamo gli sviluppi teorici indotti da anomalie come il GU, notiamo che l'impegno maggiore è confluito nel tentativo di modellizzare funzioni di utilità non-standard, nelle quali le preferenze individuali sono etero-rivolte piuttosto che strettamente egoistiche come nel modello classico.<sup>2</sup>

Il GU ha il merito di rendere particolarmente evidenti le motivazioni rivolte agli altri che influenzano il comportamento. Se queste motivazioni possano essere incorporate nel quadro della teoria dei giochi oppure no è una domanda diversa, che non affronterò in questo articolo.<sup>3</sup> Ma rispondere a questa domanda non era essenziale per il successo del GU, che fu ottenuto comunque in modo indipendente.

#### 4. ISTITUZIONI

Le prime varianti sul tema del GU erano mirate a controllare la robustezza dell'anomalia, e spesso erano esplicitamente disegnate per farla scomparire. Col senno di poi, possiamo dire che questi tentativi non hanno avuto successo. Lo status del GU non è stato scalfito, come documentano tutte le rassegne pubblicate fin dai tardi anni '80. Nei primi anni '90 tuttavia la storia del GU prende una nuova piega, con la pubblicazione di un importante articolo da parte di un gruppo internazionale di teorici dei giochi guidato da Alvin Roth.

---

<sup>2</sup> I contributi più noti sono Rabin (1993) Fehr e Schmidt (2000) Bolton e Ockenfels (2000). L'evidenza sperimentale indica in modo chiaro che i modelli puramente consequenzialisti come quello di Fehr-Schmidt e quello di Bolton-Ockenfels sono inadeguati, mentre modelli più complessi come quello di Rabin funzionano meglio, pagando però un costo di indeterminazione predittiva in un vasto insieme di giochi.

<sup>3</sup> Vedi Guala (2006).



Roth, Prasnikar, Okuna-Fujiwara e Zamir (1991) intendevano controllare due ipotesi: (1) la concorrenza è importante?, e (2) ci sono differenze rilevanti nella contrattazione in culture diverse? Per controllare la prima ipotesi essi disegnarono, accanto al GU classico, un mercato nel quale numerosi compratori competono per acquisire un bene posseduto da un singolo venditore. Sia il GU che il gioco di mercato hanno equilibri estremamente asimmetrici (nei quali un giocatore si aggiudica quasi tutta la torta). Però l'equilibrio egoista si ottiene soltanto nelle condizioni di mercato. La concorrenza ha l'effetto di "spazzare via" le considerazioni di equità.

L'esperimento di Roth è uno degli articoli più citati nell'economia sperimentale. Incorpora in un solo articolo quelli che sono generalmente considerati i tre più importanti risultati dell'economia sperimentale: la scoperta che (1) il modello dell'*homo oeconomicus* egoista non riesce a spiegare un gran numero di osservazioni; che (2) il modello della razionalità egoista può nondimeno rendere conto del comportamento in un ampio raggio di situazioni sperimentali; e infine, combinando questi due risultati, che (3) "le istituzioni sono importanti".

Vale la pena distinguere fra due tipi di istituzioni: *norme* informali come quelle di giustizia, uguaglianza, cooperazione, e reciprocità che governano il nostro comportamento in molte circostanze (come il GU); e *regole* esplicite di scambio, aggregazione, trasmissione dell'informazione, ecc. che sono tipiche di istituzioni di mercato relativamente formalizzate. L'importanza delle regole formali di scambio emerge nell'articolo di Roth dal confronto fra la contrattazione nel GU e la concorrenza di mercato. L'importanza delle norme culturali invece è messa in risalto dalle repliche dell'esperimento condotte in quattro diversi paesi (USA, Giappone, Slovenia, e Israele). Roth e colleghi trovarono delle differenze di comportamento significative: i giapponesi e gli israeliani tendono a offrire di meno (l'offerta modale è al 40%) degli americani e degli sloveni, ma le offerte inique a Tokio e Gerusalemme sono rigettate con la stessa probabilità delle offerte 50/50 di Pittsburgh e Lubiana. Nel testare un'ipotesi teorica, per la prima volta il GU veniva anche usato come uno *strumento di misurazione*.

## 5. MISURAZIONE

Nel 2000 l'antropologo Joe Henrich ha riportato una serie di osservazioni "sorprendenti" ottenute usando il GU fra i Machiguenga, un gruppo di indigeni peruviani. I Machiguenga offrivano divisioni della torta *più inique*, e rifiutavano le offerte inique *meno* spesso dei soggetti occidentali. In altri termini: si comportavano più spesso (anche se non sempre) come degli "uomini economici". I risultati dei Machiguenga potevano apparire anomali soltanto su uno sfondo di conoscenza sperimentale data per acquisita, cioè dopo che i risultati del GU nelle società occidentali erano diventati regolarità "standardizzate" e generalmente accettate. Nei primi anni '80 il comportamento dei Machiguenga sarebbe stato considerato probabilmente poco sorprendente e molto *meno* anomalo (in quanto più vicino alla previsione teorica) di quanto non appaia ora.

Henrich concluse che popoli diversi hanno diverse aspettative culturali e norme di giustizia, e che dobbiamo studiare come queste norme vengono create e sostenute in ciascun contesto sociale. Con questa idea, un gruppo di economisti e antropologi si è proposto di confrontare il comportamento in alcuni esperimenti classici (specialmente il GU e il gioco dei beni pubblici) fra quindici "micro-società" in Sud America, Africa, e Asia. Questo progetto, finanziato dalla MacArthur Foundation, è il tentativo più ambizioso ed esaltante di usare l'economia sperimentale a scopi di misurazione che sia mai stato tentato fino a oggi (vedi Henrich et al. 2004).



Il progetto MacArthur ha diverse caratteristiche metodologiche interessanti. La sperimentazione si combina con un profondo studio etnografico del contesto sociale (strutture politiche ed economiche, credenze religiose, riti, pratiche di divisione del cibo, ecc.). Queste informazioni illuminano alcuni risultati “sorprendenti” (deviazioni dalle previsioni standard) in un modo che è inusuale per le indagini sperimentali classiche. Ma soprattutto, esse sottolineano l’importanza della standardizzazione per il consolidamento dei disegni sperimentali paradigmatici, che possono essere utilizzati come strumenti portatili per misurare fenomeni sperimentali in situazioni diverse.

## 6. STANDARDIZZAZIONE

La standardizzazione è il processo di consolidamento di un disegno sperimentale paradigmatico. La standardizzazione è determinata da diversi fattori, che possono avere un peso diverso nei vari stadi di un programma di ricerca.

- (1) *Imitazione*: come ha enfatizzato Thomas Kuhn, le tecniche di una disciplina scientifica vengono spesso apprese replicando gli esemplari della generazione di scienziati precedente.
- (2) *Test di robustezza*: la logica stessa della variazione controllata richiede che i nuovi disegni siano confrontati con una base. Quest’ultima può essere una previsione teorica, ma spesso è semplicemente un risultato sperimentale “classico” che di conseguenza continua a essere replicato indefinitamente.
- (3) *Coesione disciplinare*: molte discipline a un certo punto accolgono e impongono delle “buone pratiche” di ricerca sperimentale. Pratiche come l’uso degli incentivi monetari, l’anonimità, la ripetizione, e il divieto di ingannare i soggetti fanno parte di questa categoria nel caso dell’economia sperimentale.<sup>4</sup>

Attraverso questi processi, la standardizzazione porta alla stabilizzazione dei fenomeni, e alla relativa stabilizzazione di un disegno sperimentale.<sup>5</sup> Una volta che un disegno è stato standardizzato, alcuni dei suoi dettagli perdono l’originaria giustificazione metodologica e sono conservati principalmente per ragioni pragmatiche: diventano dei mezzi per ottenere la confrontabilità fra diverse popolazioni di soggetti.

## 7. VALIDITÀ ESTERNA

Si dice spesso che affrontiamo un GU ogni volta che entriamo in un grande magazzino. Ci viene offerto un prezzo, senza l’opportunità di contrattare. Eppure in un’economia di mercato non percepiamo quasi mai i prezzi come delle offerte del tipo “prendere o lasciare”. Abbiamo solitamente altre opzioni, come per esempio entrare in un negozio che vende beni simili e controllare se ci offre dei prezzi migliori. Quando accettiamo un prezzo senza arrabbiarci, è perché crediamo (o assumiamo) che il prezzo risulti da un processo equo di concorrenza di mercato, e che sia più o meno il meglio che ci possa essere offerto data la produzione e altri

---

<sup>4</sup> Alcune di queste pratiche sono formalizzate nei cosiddetti “precetti dell’economia sperimentale” (Smith 1982) mentre altri sono giustificati in modo informale. Guala (2005, Cap. 11) include un’illustrazione e discussione più dettagliata.

<sup>5</sup> Sugden (2005) cattura entrambi questi aspetti in modo elegante per mezzo del concetto di “esemplare sperimentale” (*experimental exhibit*): un fenomeno accoppiato a un disegno standard, come il modellino di una nave in una bottiglia. Boumans (2005) sottolinea correttamente l’altro aspetto della medaglia: talvolta la robustezza di un fenomeno porta alla standardizzazione del disegno – il disegno che meglio cattura il fenomeno nella sua manifestazione “pura”.



costi. (Ovviamente questo non è sempre vero, e, infatti, le associazioni dei consumatori, gli osservatori dei prezzi, e i boicottaggi dei consumatori hanno un ruolo importante nelle economie di mercato.)

Considerazioni di equità emergono in modo diverso in diversi contesti, e il GU può al massimo essere visto come un esemplare rappresentativo di un ampio raggio di situazioni nelle quali le norme di equità giocano un ruolo più o meno diretto nel comportamento economico e sociale. Questo è particolarmente rilevante per il progetto della MacArthur Foundation. Alcune delle piccole società avevano scarsa esperienza delle transazioni di mercato, alle volte erano perfino poco abituate all'uso del denaro. Quando i disegni sperimentali facevano scattare delle norme locali era più per caso che per intenzione dello sperimentatore.<sup>6</sup>

La validità esterna è il problema di generalizzare i risultati sperimentali dalle condizioni di laboratorio ad altre situazioni che ci interessano. È solitamente considerato il “tallone di Achille” della sperimentazione nelle scienze sociali, anche se la sua rilevanza per gli esperimenti nelle scienze naturali è una questione per lo più inesplorata – e forse sottovalutata. Ho sostenuto altrove (Guala 2005, Cap. 7-11) che la questione della validità esterna deve essere risolta caso per caso. Non ha senso dubitare della validità esterna di un esperimento *in generale*. Bisogna prima chiedersi quale sia l'obiettivo dell'esperimento – quale sistema non-sperimentale si intendeva studiare in origine. Spesso gli esperimenti sono fatti senza un obiettivo esterno in mente, e dunque la validità esterna non è il loro principale scopo o preoccupazione.

Intuitivamente la standardizzazione sembra essere nemica della validità esterna. Se il mondo economico e sociale è variegato e complesso, abbiamo bisogno di flessibilità invece che di rigidità. Dobbiamo ritagliare i nostri disegni sperimentali intorno alle specifiche domande poste dal problema che stiamo studiando, piuttosto che il contrario. Ma allora perché la consolidazione di un formato rigido per il GU è stata utile e ne spiega il successo? Perché, come vedremo, i criteri metodologici che si applicano alla misurazione sono diversi da quelli che si applicano al controllo delle teorie.

## 8. OSSERVARE E INTERVENIRE

I filosofi “neo-sperimentalisti” hanno identificato l'intervento materiale come la *differentia specifica* fra la scienza sperimentale e non-sperimentale. Secondo una versione di questo approccio, sviluppata da Jim Woodward (2003) l'intervento è funzionale alla scoperta di *relazioni causali robuste*. Un intervento, secondo Woodward, è una manipolazione “chirurgica” di una variabile che lascia il resto del sistema sperimentale intatto. Consideriamo la Figura 1. Supponiamo di volere sapere se X causa Y, Y causa X, oppure se le due variabili sono correlate in modo spurio. In un intervento “chirurgico” (I) X è manipolata in modo tale da tagliare tutti i collegamenti con i suoi antecedenti causali, ma nessuno dei suoi effetti putativi è direttamente influenzato dalla manipolazione. Qualsiasi mutamento di Y dovuto a I deve passare attraverso X, in altre parole. Se l'intervento ha successo, e X causa Y, allora la variazione in X si rifletterà in una variazione in Y. Se al contrario Y causa X (il caso di mezzo nella Figura 1) oppure le due variabili sono associate soltanto in modo spurio (caso a destra) la correlazione fra X e Y sarà interrotta.

---

<sup>6</sup> Per un esempio evidente, cfr. Ensminger (2004) sull'istituzione dell' “*harambee*” fra il popolo Orma del Kenia.

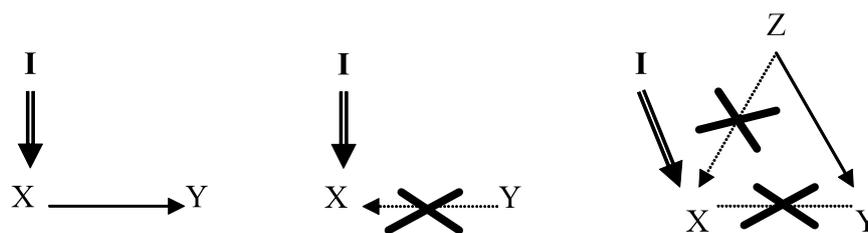


Figura 1

Usando il GU come uno strumento di misurazione, non confrontiamo due (o più) stati di una variabile che possiamo fissare a piacimento. Sicuramente stiamo confrontando *qualcosa* – il numero di rifiuti di offerte inique nelle società A e B, per esempio – ma si assume che il valore della variabile sia determinato da qualche fattore nascosto che non possiamo controllare direttamente. In effetti è estremamente importante che le procedure sperimentali non interferiscano con il funzionamento di questi fattori nascosti, perché altrimenti la misurazione sarebbe considerata “artificiale” e inutile.

Quando il GU è usato come strumento di misurazione, ovviamente, in un certo senso stiamo “controllando” qualcosa. Stiamo controllando l’ipotesi che due insiemi di osservazioni (raccolte negli Stati Uniti e fra i Machiguenga, per esempio) provengano dalla stessa popolazione o processo di generazione dei dati. È proprio per controllare questa proposizione che dobbiamo mantenere il disegno sperimentale il più stabile (o “standardizzato”) possibile. Altrimenti, non sarà possibile attribuire differenze nei comportamenti osservati a delle variazioni in qualche cosa di *diverso* (qualunque cosa esso sia) dal disegno sperimentale stesso.

Un’analogia con la termometria può essere utile: immaginiamo di volere dimostrare che l’acqua bolle a temperature diverse al livello del mare e su un altipiano. Abbiamo bisogno di un termometro standardizzato e portatile, che sia robusto a variazioni nell’ambiente esterno quando è trasportato da un sito di misurazione a un altro. Poiché non vogliamo interferire con la variabile che stiamo misurando (la temperatura) saranno permesse soltanto manipolazioni che sono funzionali all’ottenimento di una misura oggettiva e standardizzata che possa essere usata per la previsione e il confronto. Storicamente, la temperatura è stata misurata per molto tempo in modo indipendente da una comprensione anche lontanamente corretta della natura del calore (Chang 2004). Il primo passo nello sviluppo della termometria è la costruzione di un *termoscopio*. Un termoscopio misura la temperatura su una scala ordinale, e può essere usato per determinare uno o più *punti fissi* – per esempio il punto di congelamento e di ebollizione dell’acqua. Una volta fissati questi punti, possiamo costruire un *termometro* che misuri la temperatura su una scala cardinale. Storicamente, un “buon” termoscopio era valutato dalla sua conformità con le sensazioni ordinarie del caldo e del freddo: si deve espandere quando sentiamo caldo e contrarre quando sentiamo freddo. Questo garantisce che stiamo misurando qualche cosa di importante (per noi) ma ovviamente non può garantire che abbiamo incontrato una relazione fisica robusta e significativa.

Lo stesso vale per gli esperimenti economici quando sono usati come strumenti di misurazione. Le cause delle variazioni comportamentali nel GU sono ancora controverse, ed è del tutto possibile che diverse cause siano operanti in diversi contesti sociali. Nonostante questo, il GU può essere usato a scopi di misurazione, se si è ottenuta una standardizzazione appropriata e l’esperimento è stato fatto con cura. In questo caso possiamo concludere che qualsiasi differenza venga trovata, essa deve essere dovuta a differenze nelle popolazioni o nei



processi di generazione dei dati. L'inferenza, per usare l'espressione introdotta da Deborah Mayo (1996) è giustificata perché è stata *controllata severamente*.

Un controllo severo produce dati di un certo tipo D1 se l'ipotesi sotto controllo è vera, e un altro tipo di dati D2 se l'ipotesi è falsa. In termini statistici più familiari, esso minimizza la probabilità di commettere errori di tipo I e II. Il controllo severo, secondo Mayo, è il tratto distintivo della sperimentazione, perché i buoni esperimenti tipicamente supportano inferenze che sono state controllate severamente dall'esperimento stesso.<sup>7</sup> La caratterizzazione di Woodward degli esperimenti volti alla scoperta causale è per lo più coerente con questa concezione. Torniamo alla Figura 1. Se l'intervento ha successo, e X causa Y (caso a sinistra) allora le variazioni in X saranno riflesse in variazioni in Y. Se invece Y causa X o le due variabili sono associate in modo spurio (al centro e a destra) la correlazione fra X e Y scomparirà. In un esperimento genuino e ben fatto, in altri termini, si osserverà un certo tipo di dati se è vero che "X causa Y", e un altro tipo di dati se non lo è.<sup>8</sup> Qualsiasi conclusione causale arriveremo a formulare, essa sarà severamente investigata o controllata secondo il criterio di Mayo.

### 9. DATI SPERIMENTALI E RACCOLTI SUL CAMPO

Un disegno standard permette alle differenze fra diverse situazioni di emergere. Importa poco quali siano queste differenze. Potrebbero essere dovute a fattori locali, non controllati (abitudini, norme, disposizioni psicologiche); esse sarebbero nondimeno interessanti, perché queste differenze sono proprio quello che stiamo cercando. Infatti una delle lezioni più importanti degli studi interculturali riguarda la difficoltà di generalizzare su tutte le varie popolazioni

Gli studi della MacArthur Foundation mostrano che un comportamento più o meno cooperativo è correlato con diversi fattori in società diverse (ricchezza, genere, età, strutture politiche, esposizione ai mercati, ecc.). Ancora più importante è la dimostrazione che tali relazioni non sarebbero *mai* state scoperte senza la conoscenza etnografica profonda fornita dagli antropologi sul campo. Nel suo studio degli Hadza, un gruppo di cacciatori-raccoglitori in Tanzania, Marlowe (2004) per esempio riporta un grado sorprendentemente basso di condivisione e cooperazione, in una società altrimenti profondamente egualitaria. La spiegazione proposta si appella a una sottile distinzione fra ciò che si vede (per esempio la cacciagione di grossa taglia) e ciò che può essere facilmente nascosto (come il denaro). Le norme degli Hadza impongono la condivisione di ciò che si vede, ma non proibiscono effettivamente il possesso e il consumo individuale di beni più piccoli. L'avarizia è tollerata quando non può essere sanzionata in modo efficace, mentre il controllo sociale si applica alla caccia grossa.

Patton (2004) riporta un livello di comportamento cooperativo più alto fra gli Achuar che fra i Quichua, due gruppi che vivono molto vicini nella foresta amazzonica ecuadoriana. Queste relazioni si osservano sia nei dati sperimentali sia nelle pratiche di condivisione della carne osservate nelle due comunità. Tuttavia, le differenze all'interno dei gruppi mostrano che la maggiore generosità degli Achuar è quasi interamente dovuta a una piccola elite di capi che la usano per scopi politici, in particolare per creare colazioni forti e stabili fra famiglie e villaggi. I Quichua, che hanno una struttura politica meno stabile e fra i quali le coalizioni tendono a essere più effimere, semplicemente non hanno prodotto una simile classe di "condivisor politici". La conoscenza della struttura di una società e delle sue norme è cruciale, perché le

<sup>7</sup> O, più precisamente, inferenze a ipotesi che sono state controllate severamente.

<sup>8</sup> Vedi anche Woodward (2000).



correlazioni “macro” sono spesso ingannevoli o inutili. Ogni caso è diverso; ogni spiegazione è “locale”.<sup>9</sup>

## 10. TERMOMETRI SOCIALI

Un termometro è l’inizio di una diagnosi, non la fine della storia. Misura un sintomo (un sintomo importante e ubiquo) non una causa. Questo non lo rende meno utile – al contrario, staremmo molto peggio senza i termometri. Come un buon termometro, il GU è diventato uno strumento portatile che possiamo trasportare di cultura in cultura per misurare un insieme di fenomeni che riteniamo siano associati alle norme sociali. Quali siano questi fenomeni è ancora controverso. A differenza della temperatura, dove il termometro può essere confrontato con le sensazioni pure del calore e del freddo, non abbiamo delle basi simili per il GU. “Sentiamo” che il GU può essere usato per misurare ciò che appare giusto o ingiusto, ma il nostro concetto di giustizia non è necessariamente una buona guida per comprendere quelli che prevalgono nelle società che stiamo studiando. La nostra concezione dell’equità può funzionare come punto di paragone, tuttavia, e abbiamo proprio bisogno di qualcosa del genere per la misurazione. Il GU è emerso attraverso un processo di selezione sociale nella teoria dei giochi sperimentale, come un disegno robusto che cattura qualche cosa di molto importante per noi. Non abbiamo nessuna garanzia che sia lo strumento di misurazione migliore per studiare ciò che è importante per chi vive in società diverse dalla nostra. Ma è un punto di partenza, e la scienza sociale ha molto bisogno di una qualche piattaforma stabile perché possa decollare la ricerca.

## BIBLIOGRAFIA

- Bolton G. e Ockenfels A. (2000) ERC: A Theory of Equity, Reciprocity and Cooperation, *American Economic Review* 90: 166-93.
- Boumans M. (2005) *How Economists Model the World into Numbers*. Londra: Routledge.
- Chang, H. (2004) *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.
- Ensminger J. (2004) Market Integration and Fairness: Evidence from Ultimatum, Dictator, and Public Goods Experiments in East Africa, in Henrich et al. (2004) 356-81.
- Fehr E. e Schmidt, K. (1999) A Theory of Fairness, Competition and Co-operation, *Quarterly Journal of Economics* 114: 817-68.
- Gil-White F. (2004) Ultimatum Game with an Ethnicity Manipulation: Results from Khodovin Bulgan Sum, Mongolia, in Henrich et al. (2004) 260-304.
- Guala F. (2005) *The Methodology of Experimental Economics*. New York: Cambridge University Press.
- Guala F. (2006) Has Game Theory Been Refuted?, *Journal of Philosophy* 103: 239-263.
- Gurven M. (2004) Does Market Exposure Affect Economic Behavior? The Ultimatum Game and the Public Goods Game among the Tsimane’ of Bolivia, in Henrich et al. (eds. 2004) 194-231.

---

<sup>9</sup> Si vedano anche Gurven (2004) Tracer (2004) e Gil-White (2004) per altri esempi del genere.



- Güth W., Schmittberger R. e Schwarz, B. (1982) An Experimental Analysis of Ultimatum Bargaining”, *Journal of Economic Behavior and Organization* 3: 367-388.
- Henrich J. (2000) Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon”, *American Economic Review* 90: 973-79.
- Henrich J., Boyd R., Bowles S., Camerer C., Fehr, E. e Gintis H. (eds. 2004) *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford: Oxford University Press.
- Marlowe F. (2004) Dictators and Ultimatums in an Egalitarian Society of Hunter-Gatherers: The Hadza of Tanzania, in Henrich et al. (2004) 168-193.
- Mayo D. (1996) *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.
- Patton J. (2004) Coalitional Effects on Reciprocal Fairness in the Ultimatum Game: A Case from the Ecuadorian Amazon, in Henrich et al. (2004) 96-124.
- Rabin M. (1993) Incorporating Fairness into Game Theory and Economics, *American Economic Review* 83: 1281-302.
- Roth A., Prasnikar V., Okuno-Fujiwara M. e Zamir, S. (1991) Bargaining and Market Behavior in Jerusalem, Lubljana, Pittsburgh and Tokyo: An Experimental Study, *American Economic Review* 81: 1068-1095.
- Smith V. (1982) Microeconomic Systems as an Experimental Science, *American Economic Review* 72: 923-55.
- Tracer D. (2004) Market Integration, Reciprocity, and Fairness in Rural Papua New Guinea: Results from a Two-Village Ultimatum Game Experiment, in Henrich et al. (2004) 232-259.
- Sugden R. (2005) Experiments as Exhibits and Experiments as Tests, *Journal of Economic Methodology* 12: 291-302.
- Woodward J. (2000) Data, Phenomena, and Reliability, *Philosophy of Science* 67: S163-179.
- Woodward J. (2003) *Making Things Happen*. New York: Oxford University Press.