

*di Stefano Nobile*

## I. Il problema dell'operativizzazione di concetti complessi

---

Di consuetudine, nella ricerca sociale di impronta quantitativa con questionario (rivolto a individui che possano e vogliano fornire delle risposte) si rilevano tre tipi di informazioni: quelle di carattere socio-anagrafico (quanti anni hanno, dove sono nati, qual è la loro condizione occupazionale e via dicendo); quelle relative a comportamenti (con quale frequenza hanno visitato una mostra d'arte nell'ultimo anno o se si sono recati alle urne in occasione delle più recenti consultazioni elettorali); infine, quelle relative a opinioni, atteggiamenti e valori<sup>1</sup> (*cfr.* cap. 4). Queste tre famiglie di informazioni si differenziano almeno sotto due aspetti: quello dell'impegno cognitivo e quello temporale. Passando dalle domande socio-anagrafiche a quelle sulle opinioni e gli atteggiamenti, infatti, lo sforzo di concentrazione che deve compiere l'intervistato è ben diverso. Altro è, infatti, ricordare quale sia il proprio titolo di studio o se si è sposati o meno, altro è offrire un'opinione sull'opportunità o meno di ripristinare l'energia nucleare in Italia o assumere una posizione rispetto alle famiglie arcobaleno. Nel caso degli atteggiamenti e delle opinioni, infatti, non è affatto raro che l'intervistato possa trovarsi per la prima volta – l'occasione dell'intervista – a dover rispondere a domande rispetto alle

---

<sup>1</sup> Si tratta, con ogni evidenza, di una semplificazione. Una domanda come “quante persone lavorano nella tua stessa stanza?”, infatti, non è riconducibile a nessuna delle tre famiglie indicate. Resta comunque vero che una parte ragguardevole dei questionari è imperniata sulla triade caratteri socio-anagrafici / comportamenti / atteggiamenti.

quali non ha mai avuto occasione di riflettere se non, magari, superficialmente.

La differenza tra le tre famiglie di domande sta anche nella dimensione temporale. Possiamo dire, con un minimo di beneficio d'inventario, che le domande di carattere socio-anagrafico riguardano prevalentemente il presente (quanti figli hai, in quale settore professionale lavori); quelle sui comportamenti riguardano invece il passato e richiedono spesso uno sforzo mnemonico all'indietro (quante abitazioni hai cambiato nel corso della tua vita? Quanti viaggi all'estero hai fatto nell'ultimo anno?); infine, quelle sugli atteggiamenti e le opinioni riguardano il futuro o, comunque, possono avere implicazioni su di esso (cosa pensi del politico x e delle promesse fatte in campagna elettorale? Qual è la tua posizione rispetto all'eutanasia?). Potrebbe non essere immediatamente intuitivo, eppure il punto di partenza per poter parlare di tecniche di rilevazione degli atteggiamenti sta proprio qui: un atteggiamento fa da innesco a un comportamento. Per quanto possa esprimere comportamenti impulsivi, un individuo prima elabora – anche in maniera del tutto superficiale – determinate credenze e poi eventualmente mette in atto i comportamenti che seguono quelle credenze. In questo senso, l'atteggiamento è una disposizione ad agire (Searle, 1983) e, dunque, in un certo modo prefigura possibili azioni future. Per molti versi, perciò, gli atteggiamenti sono uno degli aspetti più interessanti da rilevare con la ricerca sociale di impronta quantitativa, proprio perché permettono di aprire possibili scenari futuri in un'ottica predittiva.

Spesso i concetti di opinione e quello di atteggiamento vengono impiegati come sinonimi. I due concetti, tuttavia, sono tenuti insieme soltanto perché sono trattabili pressoché nello stesso modo nella ricerca sociale, attraverso le tecniche di scaling. Ma va chiarito che l'*atteggiamento* rappresenta «quell'insieme di tendenze e sentimenti, pregiudizi e nozioni preconcepite, idee, timori, apprensioni e convinzioni di una persona nei confronti di un particolare argomento» (Thurstone e Chave 1929; trad. it. 1974, p. 97); mentre con *opinione* ci si riferisce «all'espressione verbale dell'atteggiamento» (ibidem). Corbetta (2014, p. 252) chiarisce che «l'atteggiamento è il concetto, le opinioni ne sono gli indicatori». Ad avviso di chi scrive, le opinioni sono qualcosa di più volatile degli atteggiamenti. Esse, infatti, possono essere fornite senza che necessariamente

l'intervistato ne possieda una radicata convinzione<sup>2</sup>. Al contrario, gli atteggiamenti sono il precipitato di un insieme di disposizioni e visioni del mondo che preparano i comportamenti. Quest'ultima caratteristica è quella che tiene insieme i due approcci principali alla definizione dell'atteggiamento: quella comportamentista e quella cognitivista. Per la prima, l'atteggiamento è l'esito del complesso di risposte che l'individuo dà attraverso meccanismi di riposta a uno stimolo e che, col tempo, si radicano nella sua visione del mondo. Per i cognitivisti, invece, l'atteggiamento fa capo a una struttura psico-cognitiva stabile. In entrambi i casi, come si diceva, è condivisa l'idea che l'atteggiamento funga da innesco potenziale all'azione. Essi sono comunque concepiti come una proprietà continua, a cui possono corrispondere incrementi differenziati in maniera infinitesimale

Poiché le opinioni e gli atteggiamenti sono concetti ad alto livello di astrazione (dal momento che possono riguardare il pregiudizio etnico, l'autoritarismo, la religiosità, l'omofobia, la prudenza, il conformismo, la superstizione, l'alienazione, la conflittualità e così via), essi sono magmatici, sfuggenti, certamente difficili da cogliere con strumenti standardizzati (sebbene con quelli non standardizzati i problemi interpretativi non siano da considerarsi affatto minori). La psicomatria e la sociologia hanno prodotto una serie di dispositivi congegnati proprio per poter trasformare gli atteggiamenti in qualcosa di misurabile. Questi strumenti si chiamano *scale di atteggiamento*. Gli atteggiamenti, al pari delle opinioni e dei tratti della personalità, non sono entità discrete, bensì continue. Vale a dire che l'intensità di un certo atteggiamento, opinione o tratto non varia "per salti", ma attraverso un ipotetico continuum che oscilla tra un minimo e un massimo di intensità. Essi, pertanto, sono misurabili nella misura in cui fanno riferimento a variazioni infinitesimali su un ipotetico continuum. Ma in che senso possiamo parlare di una *misurazione* degli atteggiamenti (Capecchi 1962)? Partiamo da un'idea un po' grossolana, ma efficace: se dicessimo che Anna è progressista 8 in una scala da 1 a 10 è come se stessimo esprimendo una sorta di "temperatura" del livello di progressismo di Anna, ben sapendo – tuttavia – che la risposta ottenuta fa riferimento a un'unità di misura che è solo formale, ma che poi viene adottata in maniera soggettiva da ciascun rispondente. L'acquisizione di questo punteggio non è frutto di un procedimento arbitrario ma, al tempo stesso,

---

<sup>2</sup> Pavsic e Pitrone (2003, p. 28) sintetizzano la questione definitoria vedendo nell'atteggiamento uno stato interiore e nell'opinione la sua manifestazione pubblica.

non è neppure così scontata e indefettibile come potrebbe sembrare. Come si è già detto, quando si vuole rilevare un atteggiamento, un'opinione o un valore ci si addentra in un territorio insidioso per varie ragioni. La prima di queste è che i concetti ad alto livello di astrazione che richiamano una disposizione (ossia qualcosa che esiste in potenza, ma che ha bisogno di essere attivato per potersi mostrare) richiedono sempre una pluralità di indicatori. Soltanto così il ricercatore può assicurarsi che il concetto in qualche modo possa essere intercettato nella maniera meno approssimativa possibile. Ciò necessita l'individuazione di indicatori che abbiano una relazione semantica<sup>3</sup> plausibile con il concetto al quale si riferiscono e la loro identificazione non sempre è così semplice.

Un secondo problema è che, nel caso specifico di opinioni e atteggiamenti, questi ultimi non necessariamente preesistono al momento della loro rilevazione in occasione della survey (Pavsic e Pitrone, 2003) e, per di più, sono instabili. Il che dà luogo a un potenziale paradosso: all'interno di un determinato campione di intervistati potrebbe esserci una minoranza di persone con un'opinione ben radicata (per esempio, una piccola quota di soggetti favorevoli all'aborto in un'indagine su questo tema) e una larga maggioranza con opinioni vacue, improvvisate, magari dichiarate emotivamente sulla scorta dell'atteggiamento rassicurante da offrire al ricercatore che non si discosti dall'ortodossia di una certa etica cattolica. È inoltre possibile, e tutt'altro che raro, che l'intervistato fornisca un'opinione sia davanti a questioni palesemente assurde (proposte con il preciso scopo di testare l'inclinazione degli intervistati a fornire *comunque* una risposta), sia davanti a modalità di risposta volutamente insensate. Un esempio icastico è quello riportato da Hartley (1946) in un suo studio sul pregiudizio. In esso, l'autore dimostrò che i tre quarti circa degli intervistati che avevano dichiarato di disprezzare gli ebrei, sostenevano di

---

<sup>3</sup> Non sempre la ricerca sociale si basa esclusivamente sulla dimensione *semantica* del rapporto di indicazione. In alcuni casi essa poggia su una mera *associazione statistica*. Significa che i ricercatori hanno trovato una certa relazione tra un indicatore e l'apparizione di un certo tratto, sintomo, caratteristica o proprietà che dir si voglia. Un esempio lampante è quello rappresentato da un test ancora oggi usatissimo per le diagnosi psichiatriche: il *Minnesota Multiphasic Personality Inventory* (MMPI-2). Esso raccoglie un mezzo migliaio di domande, tra le quali questa: "sento prurito alle dita quando apro la maniglia di una porta". Se l'affermazione riceve una risposta affermativa dal soggetto sottoposto a test, per gli autori (Hathaway e McKinley, 1939) indicherebbe una relazione di causa-effetto con alcune importanti proprietà psichiche. Ma, come precisa Marradi (2007, p. 170), «questi sono casi-limite, che si possono giustificare solo se si ha bisogno di indicatori di un concetto X e non si trovano concetti la cui associazione empirica col concetto X si possa giustificare semanticamente».

detestare anche altri popoli come i valoniani, i pirenesi e i danireici. È superfluo dire che nessuno di questi tre popoli esiste né è mai esistito.

Opinioni e atteggiamenti, dunque, vanno maneggiati con estremo tatto, sia perché – come si è visto – sono esposti a possibili risposte bizzarre, sia perché, da un punto di vista tecnico-visuale, si presentano con delle caratteristiche che possono affievolire l'attenzione dell'intervistato, dando luogo a un potenziale effetto noto come *response set*<sup>4</sup>. A differenza delle domande di cui si è prevalentemente parlato nel capitolo 4, infatti, le scale di atteggiamento si presentano con una veste particolare (ma non esclusiva), ossia in forma di batteria. Quest'ultima è una sequenza di item<sup>5</sup> rispetto ai quali l'intervistato deve rispondere avendo a disposizione sempre le stesse opzioni: sì/no<sup>6</sup>, oppure "completamente in disaccordo", "abbastanza in disaccordo", "né d'accordo né in disaccordo", "abbastanza d'accordo" e "completamente d'accordo", o "vero", "falso", "non so", o – ancora – può esprimere la sua opinione su una scala di valori numerici ai cui estremi sono posti quello massimo e quello minimo.

L'insieme delle risposte date agli item fornisce un punteggio complessivo – al netto di alcuni accorgimenti tecnici legati all'orientamento positivo negativo degli item – sia rispetto all'oggetto della scala (per esempio, la fiducia nei confronti delle istituzioni), sia rispetto ai singoli intervistati. In genere, la dimensione sottostante le scale di atteggiamento mira a collocare i soggetti lungo un continuum che si muove tra una posizione favorevole e una sfavorevole, tra una migliore e una peggiore, tra una preferita e una non preferita e così via. Tuttavia, non sempre le scale intercettano un'unica dimensione latente (che poi dovrebbe essere quella del corrispondente concetto sul piano teorico). Alcune di esse (come il differenziale semantico; Osgood, Suci, Tannenbaum 1957), al contrario, mirano a

---

<sup>4</sup> Il *response set* è l'attitudine che l'intervistato *potrebbe* manifestare dando risposte sempre uguali a elementi della batteria presentati in sequenza: per esempio, rispondendo sempre "sì", o "abbastanza d'accordo" anche a stimoli che hanno verso contrario (favorevoli o sfavorevoli rispetto a un certo tema). Si tornerà più dettagliatamente su questo tema nelle pagine che seguono.

<sup>5</sup> Le scale presentano un'unica domanda che chiarisce all'intervistato, attraverso opportune istruzioni, come egli debba rispondere. Le voci che compongono la batteria prendono il nome di *item*, un termine anglosassone che sta a indicare un elenco.

<sup>6</sup> Thurstone e Chave (1929), per esempio, studiarono l'atteggiamento di un campione di statunitensi nei confronti della Chiesa. Per farlo, ricorsero a trecento persone alle quali fu chiesto, nella fase di preparazione della scala, di collocare l'intensità dell'atteggiamento espresso da 130 item su un continuum tra 1 e 11, per poi selezionare quelli che esprimevano valutazioni più omogenee e sottoporli a un campione di intervistati. Al livello minimo della scala si collocavano frasi come "la Chiesa è un parassita della società", mentre sul versante opposto si trovavano frasi come "io credo che la Chiesa sia la più grande istituzione in America".

evidenziare più dimensioni: il che ci induce – come si vedrà meglio in seguito – a distinguere tra scale monodimensionali e multidimensionali<sup>7</sup>.

Prima di passare a una rassegna delle diverse tecniche di scaling, è opportuno qualche sommaria indicazione su alcuni degli elementi che caratterizzano la struttura delle scale di atteggiamento nonché un brevisimo excursus storico.

Un primo problema da affrontare riguarda la monodimensionalità (o unidimensionalità) della scala. Poiché questo strumento è concepito per rilevare atteggiamenti che sono immaginati come continui e i cui indicatori rilevano tutti la stessa proprietà, si suppone che ogni item che compone la scala rimandi allo stesso concetto di partenza, il che rimanda tanto a un problema in termini di rapporto di indicazione (*cfr.* Cap. 3) quanto a un problema di validità (*cfr.* Cap.10).

Un altro tratto da considerare è l'autonomia semantica delle modalità di risposta delle affermazioni che compongono la scala. Con l'espressione autonomia semantica si intende che ciascuna delle risposte offerte all'intervistato non dipende dalle altre, in quanto il suo significato risulta essere chiaro e autonomo. La sua interpretazione, dunque, non necessita né della conoscenza delle altre modalità di risposta, né dell'intestazione della domanda per poter essere decodificata. Se la domanda fosse *In quale regione risiedi?* l'intervistato non avrebbe bisogno di passare in rassegna tutte e venti le regioni italiane, ma andrebbe a cercare direttamente la risposta nella lista, Basilicata, Lazio o Campania che sia. Quando invece all'intervistato – come accade per esempio in una scala come la Likert – le alternative sono “del tutto in disaccordo, piuttosto in disaccordo, né d'accordo, né in disaccordo, piuttosto d'accordo e del tutto d'accordo, allora siamo di fronte a risposte a parziale autonomia semantica: l'intervistato ha bisogno di conoscere almeno i poli del continuum lungo il quale deve esprimere la propria risposta. Le risposte *sì* e *no*, al contrario, non hanno alcuna autonomia semantica: esse sono regolate dall'eventuale presenza di gradienti di scala (*sì*, ma...; *no*, ma...), quanto dall'ancoraggio all'intestazione della domanda. Quanto a quest'ultimo aspetto, le scale sono dette autoancoranti quando vengono fissati i poli estremi della scala, ma la zona intermedia può essere caratterizzata da caselle, numeri naturali, segmenti, simboli. Nel caso delle scale autoancoranti – come vedremo più

---

<sup>7</sup> Quelli presentati non sono i soli modi per classificare le scale di atteggiamento. Goode & Hatt (1968), per esempio, le distinguono tra scale «che misurano la distanza sociale, scale di valutazione, scale di graduazione, scale basate sulla coerenza interna e scale basate sull'analisi della struttura latente» (*ibi*, p. 376).

avanti – viene data al rispondente una delega rispetto al posizionamento sul continuum, che non è fissato – come negli altri tipi di scala – sui diversi picchetti che lo attraversano.

Quanto alla storia della tecnica delle scale, l'origine delle tecniche di scaling può essere fatta risalire alla seconda metà dell'Ottocento. A quell'epoca Francis Galton – cugino di Darwin e sostenitore di pericolose idee eugenetiche – cercò di individuare dei criteri per rilevare il concetto di intelligenza. Tra questi criteri ve n'erano alcuni decisamente discutibili, come la prontezza di reazione a un segnale acustico. Fu sulla base dei tentativi di Galton di misurare l'intelligenza che il francese Alfred Binet, nel 1905, riuscì a varare quello che possiamo considerare come il primo strumento di scaling: un test per la misurazione dell'età mentale, legato a prove di difficoltà crescente. L'idea di fondo, quella secondo la quale le risposte a una serie di item potesse essere uno strumento idoneo a rilevare proprietà ad alto livello di astrazione, venne conservata e utilizzata per strumenti mirati a rilevare non solo capacità ma anche, appunto, atteggiamenti. La tecnica delle scale diede in seguito contributi fondamentali tra gli anni Venti e gli anni Quaranta del secolo scorso: dapprima con Bogardus (1925), Thurstone (1928), Likert (1932) quindi con Guttman (1944) e Cantril (1946). Nel corso del tempo a queste scale se ne sono aggiunte altre, che hanno cercato di risolvere i problemi tecnici e semantici delle precedenti. Fatto sta che a tutt'oggi la scala proposta da Likert – una delle più semplici da usare – è quella che continua a campeggiare con maggiore frequenza su molti questionari, dando luogo a un uso spesso indiscriminato (Marradi e Gasperoni 2002). Come il lettore avrà intuito, le tecniche di scaling non sono tutte uguali e, di conseguenza, non possono essere impiegate indistintamente, secondo l'uzzolo del ricercatore. Al tempo stesso, esse rispondono a obiettivi che suggeriscono una possibile classificazione: scale centrate sui soggetti, sugli stimoli e sulle risposte (Giampaglia, 1990, pp. 42-43). Quelle centrate sui soggetti, come le scale Likert, presuppongono che la variazione delle risposte agli stimoli sia attribuibile alle differenze individuali. Le scale centrate sugli stimoli si basano invece sull'idea che «la variazione sistematica nelle reazioni dei soggetti agli stimoli» (Giampaglia, 1990, p. 42) debba essere attribuita «alle differenze tra gli stimoli rispetto a una certa proprietà» (*ibidem*). Un buon esempio ne è la scala Thurstone. L'approccio centrato sulle risposte, infine, contempla quelle scale nelle quali la variazione delle reazioni dei soggetti è attribuita tanto alla variazione dell'intensità con cui i soggetti posseggono una certa

proprietà, quanto alla variazione dell'intensità dello stimolo. Sono le scale più difficili da progettare – va chiarito a chi si accosta a questi strumenti per la prima volta – e da elaborare in sede di analisi dei dati. Esse si distinguono, a loro volta, tra modelli deterministici (come, per esempio, la scala Guttman) e modelli probabilistici (come la scala Rasch). A quest'ultima famiglia saranno fatti soltanto alcuni cenni. Di seguito, invece, saranno passate in rassegna le tecniche di scaling di maggior uso.

## 2. La scala di distanza sociale di Bogardus

---

Negli anni Venti del Novecento un sociologo formatosi nella prestigiosa Scuola di Chicago, Emory S. Bogardus, varò quella che viene considerata come la prima scala d'atteggiamento nell'ambito delle scienze sociali, al di là dell'opera pionieristica di Binet. Questa tecnica di scaling avrebbe in seguito campeggiato sui manuali con il nome di scala di distanza sociale di Bogardus. Essa aveva l'obiettivo di "misurare" il pregiudizio etnico, un tema assai sentito nella comunità sociologica di Chicago<sup>8</sup>, per via delle rapide trasformazioni alle quali la società americana stava andando incontro a seguito del forte flusso migratorio. Attraverso questo strumento, agli intervistati veniva domandato cosa fossero disposti ad accettare rispetto a una serie di popolazioni di origine nazionale diversa: inglesi, svedesi, polacchi e coreani. La scala prevedeva una serie di item di intensità crescente (una netta anticipazione del modello deterministico di Guttman, sbandierato dallo stesso psicometrico come la prima scala ad item cumulativi). Si domandava cioè agli intervistati se fossero inclini a escludere un inglese (o uno svedese, o un polacco o un coreano) dal proprio Paese; se fossero disposti ad accettarlo solo come semplice turista; se lo avessero accolto come cittadino della sua stessa nazione; se lo avessero accettato come collega di lavoro; se avrebbero accettato di averlo come vicino di casa; se lo avrebbero potuto includere nella loro cerchia di amici e, infine, se avessero acconsentito ad avere con lui una qualche forma di parentela. La scala – che tra il 1926 e il 1967 fu testata dallo stesso

---

<sup>8</sup> Non a caso, a introdurre il concetto di *distanza sociale* – da intendersi come un continuum che va da un massimo di intimità a un massimo di ostilità – fu Robert E. Park (1902), uno degli esponenti di punta della Scuola di Chicago.

Bogardus e dai suoi allievi su oltre 8.000 soggetti – dava luogo a due distinti punteggi complessivi: l'indice di contatto sociale (SCR, ossia *Social Contact Range*) e l'indice di distanza sociale (SCD, ovvero *Social Contact Distance*). Come si applica la scala? Si attribuisce un punto per ogni risposta affermativa, zero per ogni risposta negativa. Tuttavia, se si guarda la Tabella 7.1 si capisce subito che qualcosa non torna. La tabella riporta la percentuale delle risposte affermative incassate complessivamente dai quattro gruppi nazionali. Le ultime due colonne fanno registrare un brusco crollo. Ciò significa che il punteggio, calcolato sia rispetto ai diversi gruppi che ai singoli intervistati, per cinque affermazioni su sette esprime una disposizione positiva e accogliente nei confronti dell'oggetto di riferimento, mentre ne esprime una negativa negli ultimi due item. Dunque, attribuire punteggio uno al primo item (espulsione), due al secondo, eccetera, fino a sette a quello sul matrimonio è scorretto. Il problema rispetto ai due item "accettarlo solo come visitatore" ed "espellerlo" può tuttavia essere risolto facilmente. In questo – come nei casi di altre scale che saranno trattate nelle prossime pagine – basta avere l'accortezza di trasformare il punteggio delle risposte positive in zero e quello delle positive in uno. In questo modo i conti tornerebbero sia nell'analisi condotta sui diversi gruppi, sia rispetto ai punteggi dei singoli intervistati. In *extrema ratio*, è possibile cassare gli ultimi due item<sup>9</sup>.

**Tabella 7.1** Valori percentuali relativi alla distanza sociale da alcune minoranze (risposte affermative). Adattamento da Bogardus (1925, p. 25). N = 1.725.

	<b>Matrimonio (7)</b>	<b>Amico (6)</b>	<b>Vicino di casa (5)</b>	<b>Collega di lavoro (4)</b>	<b>Cittadino (3)</b>	<b>Solo come visitatore</b>	<b>Espulsione (1)</b>
Inglesi	93,7	96,7	97,3	95,4	95,9	1,7	0,0
Svedesi	45,3	62,1	75,6	78,0	86,3	5,4	1,0
Polacchi	11,0	11,6	28,3	44,3	58,3	19,7	4,7
Coreani	1,1	6,8	13,0	21,4	23,7	47,1	19,1

<sup>9</sup> Il punteggio sui singoli intervistati si otterrà come somma dei prodotti tra gli item che hanno ricevuto risposte affermative (negative per gli ultimi due) e il rispettivo punteggio di ponderazione (da sette a uno). Per i gruppi nazionali, il punteggio si ottiene tramite la somma dei prodotti tra percentuali di risposte affermative e punteggio di ponderazione. I coreani, ad esempio, avranno come punteggio  $(1,1 \times 7) + (6,8 \times 6) + (13 \times 5) + (21,4 \times 4) + (23,7 \times 3) + [(100 - 47,1) \times 2] + [(100 - 19,1) \times 1]$ . Gli ultimi due item sono stati calcolati come complemento a 100 delle risposte affermative ricevute.

Se, inoltre, si assume che gli item della scala siano equidistanti tra loro, lo strumento mostra ulteriori pecche. Infatti, nulla ci assicura che la distanza, per esempio, tra accettare un coreano come collega di lavoro ed accettarlo come cittadino sia la stessa che lo separa dall'accettarlo come vicino di casa. In più, come hanno fatto notare Goode e Hatt (1968, p. 380), «non si conosce un punto zero. Di conseguenza è impossibile concludere che gli svedesi, ad esempio, siano accettabili approssimativamente il doppio dei polacchi, i quali sono a loro volta il doppio dei coreani» (cfr. Tabella 7.2).

**Tabella 7.2** Distanza sociale di quattro minoranze etniche

	<b>Matrimonio</b>	<b>Amico</b>	<b>Vicino di casa</b>	<b>Collega di lavoro</b>	<b>Cittadino</b>	<b>Solo come visitatore</b>	<b>Espulsione</b>	<b>Totale</b>
Inglesi	655,9	580,2	486,5	381,6	287,7	196,6	100,0	2688,5
Svedesi	317,1	372,6	378	312	258,9	189,2	99,0	1926,8
Polacchi	77,0	69,6	141,5	177,2	174,9	160,6	95,3	896,1
Coreani	7,7	40,8	65,0	85,6	71,1	105,8	80,9	456,9

Proposta così, sembrerebbe che la scala Bogardus abbia più difetti che pregi. Ma non è così. Va infatti sottolineato che la scala è ancora molto diffusa perché può essere applicata a valori, gruppi sociali, individui<sup>10</sup>. Per esempio, in una ricerca sulla xenofobia la si potrebbe applicare a gruppi sociali oggetti di pregiudizio, come omosessuali, tossicodipendenti, immigrati, eccetera, avendo l'accortezza di calibrare gli item sull'oggetto di riferimento.

### 3. La scala di Thurstone

L'idea di poter costruire una scala unidimensionale che presenti degli intervalli che sembrano uguali (di cui si sono già discussi i limiti nella

<sup>10</sup> Per una panoramica sulle potenziali applicazioni più recenti della scala Bogardus, cfr. Caselli (2008).

scala Bogardus) illuminò i lavori dello psicometrico Louis Thurstone. Il suo lavoro partiva dagli studi di psicometria – molto in voga in quegli anni – relativi allo studio delle differenze percettive, come i ben noti lavori di Fechner. Negli anni Venti Thurstone pubblicò dunque tre articoli nei quali propose altrettante scale di atteggiamento. La spuntò quella ad item che sembrano uguali in versioni che, nell'assegnazione dei punteggi ai soggetti, ha due varianti. Le vedremo tra breve.

La scala consiste nella messa a punto di undici item, che variano tra 0 e 10<sup>11</sup>, di cui si presuppone la capacità di rilevare l'atteggiamento rispetto all'oggetto di analisi in forma crescente. Col suo collega Ernest John Chave, Thurstone intendeva rilevare gli atteggiamenti di un campione di intervistati nei confronti della Chiesa (Thurstone e Chave, 1929). Nelle parole degli stessi autori, «ad alcuni gruppi e a molti individui fu chiesto di scrivere la loro personale opinione verso la Chiesa; venne inoltre fatta un'opera di ricerca nella letteratura corrente per trovare brevi affermazioni adeguate che potessero servire per la scala<sup>12</sup>. Dal materiale così raccolto si formò una lista di 130 affermazioni che erano l'espressione di atteggiamenti che, nella misura in cui era possibile, coprivano tutta la gradazione della scala, da un estremo all'altro» (*ivi*, p. 111). I 130 item dovevano essere sottoposti al vaglio di alcuni "giudici", ossia degli esperti che avevano il compito di collocare ciascun item in uno degli undici picchetti di scala<sup>13</sup>. Gli undici item che vengono selezionati dovranno avere due proprietà: il loro punteggio sarà pari alla mediana dei punteggi assegnati dai giudici; a parità di punteggi, viene selezionato l'item che mostra, per ciascun picchetto di scala, la minore variabilità. Va da sé che, in teoria, quanto più numerosi sono i giudici, tanto più è accurato il processo di selezione.

Vediamo, dunque, come avviene questa selezione. Immaginiamo di avere a disposizione cinque item che sono tutti in concorrenza tra loro per occupare il quarto picchetto di scala. I giudici chiamati a valutarli sono sette, come da Tabella 7.4.

---

<sup>11</sup> La scelta di far variare i punteggi degli item – che, si badi bene, *non* sono noti agli intervistati – è in parte discutibile, poiché presume il 5 come punto centrale di scala. Quest'ultima sarebbe più intellegibile se il campo di oscillazione avesse un'escursione tra -5 e +5, con 0 come punto di equilibrio centrale. In ogni caso, il problema non si pone per gli intervistati.

<sup>12</sup> Questa selezione, per così dire "a spanne" del materiale empirico, ha condotto osservatori come Cannavò (2003, p. 30) a parlare di «induttivismo assoluto» insito nella scala Thurstone.

<sup>13</sup> I picchetti sono i punti di ancoraggio delle scale di atteggiamento. Nel caso della scala Thurstone sono i singoli valori di scala (da 0 a 10 o da 1 a 11); in quello delle scale Likert sono le risposte a parziale autonomia semantica.

**Tabella 7.3** Punteggi assegnati da sette giudici a cinque item che hanno la stessa mediana

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>
G1	5	5	4	3	3
G2	2	3	5	6	7
G3	5	6	3	5	5
G4	5	2	6	3	3
G5	1	2	3	4	4
G6	3	8	4	4	1
G7	4	4	1	9	2
Deviazione standard	1,62	2,21	1,60	2,12	1,99
Mediana	4	4	4	4	3

I cinque item hanno tutti lo stesso valore mediano (4). Quale scegliere? Quello che ha la più bassa deviazione standard, che indica un più alto grado di accordo nei punteggi assegnati dai giudici. Dunque, nella scala finale entrerà l'item C.

Rimane da dire come si debbano assegnare i punteggi agli intervistati. È qui che entrano in gioco le due varianti proposte da Thurstone: nella prima, gli intervistati – ai quali gli item della scala vengono proposti in maniera accidentale, in modo che non possano percepire la progressione da quello più ostile a quello più favorevole all'oggetto di atteggiamento – devono rispondere se sono d'accordo o in disaccordo con ciascuno degli undici item finali. A quel punto al singolo intervistato viene attribuito un punteggio pari alla media (o alla mediana, una variante nella variante) dei punteggi corrispondenti agli item con i quali egli si è detto d'accordo. Questa soluzione, peraltro la più praticata, espone però la scala a un serio rischio di *response set* (Marradi, 1981). L'alternativa è quella di far leggere all'intervistato tutti e undici gli item e poi chiedere di scegliere quello che più si avvicina al suo punto di vista. In questo secondo caso, il punteggio attribuito al singolo rispondente equivale al valore di scala dell'item scelto. Ma anche questa seconda variante presenta degli inconvenienti. Prendiamo a titolo di esempio una porzione degli item proposti da Thurstone e Chave:

1. Mi pare che l'insegnamento della Chiesa sia nell'insieme troppo superficiale per essere socialmente significativo.
2. Credo che la Chiesa soddisfi abbastanza gli interessi dei giovani.

3. Sento che le funzioni mi ispirano e mi aiutano a vivere e ad essere migliore per tutta la settimana.
4. Mi sembra che la Chiesa mantenga la vita economica e politica a un livello più elevato di quanto non sarebbe altrimenti.
5. Trovo che le funzioni della Chiesa danno quiete e ispirazione.
6. Quando vado in Chiesa godo delle funzioni liturgiche con buona musica.

Il secondo item, per esempio, non necessariamente esprime un atteggiamento di orientamento positivo nei confronti della Chiesa, considerato nella sua globalità. L'item, infatti, potrebbe venire selezionato anche da un non credente (Cannavò, 2003, p. 30).

A titolo di ulteriore esempio, di seguito viene riportata anche una scala prodotta da Pavsic e Pitrone (2003, p. 57), che dà conto della crescente intensità degli item rispetto al tema della tossicodipendenza.

1. Ogni volta che sento che un giovane è morto per overdose provo una sensazione di sollievo.
2. Non ho alcuna comprensione verso i tossicodipendenti.
3. Il tossicodipendente finisce sempre per diventare uno spacciatore.
4. Per colpa dei tossicodipendenti camminare per strada è diventato pericoloso.
5. I nostri giovani si drogano perché sono stati abituati ad avere tutto.
6. Nessuno di noi può illudersi che il problema della tossicodipendenza non lo tocchi.
7. Il giovane che si droga è un disperato.
8. Credo che il dramma dei drogati sia un vero affare per tanta gente insospettabile.
9. Contro il tossicodipendente non servono misure repressive, ma una seria prevenzione.
10. Un giovane drogato ha bisogno di rispetto e di solidarietà.
11. Ogni volta che vedo un giovane distrutto dalla droga non posso fare a meno di sentirmi responsabile.

Poiché presenta un certo grado di macchinosità, a cominciare dalla necessità di assoldare degli esperti per la valutazione degli item, la scala di Thurstone è andata progressivamente in disuso. Ciò nonostante, essa a tutt'oggi rappresenta un nobile tentativo di costruire uno strumento che ha lo scopo di graduare l'intensità degli item. In più, non va sottovalutato il fatto che, generalmente, questa scala manifesta un livello di affidabilità generalmente soddisfacente (Perrone, 1982, p. 118), quantunque,

ovviamente, questa affermazione non possa essere fatta in astratto, ma solo con riferimento alle effettive capacità di progettazione dei ricercatori<sup>14</sup>.

## 4. La scala di Likert

Negli anni Trenta, Rensis Likert (1932) – uno studioso con una formazione da economista, sociologo e psicologo – progettò una scala di atteggiamento che mirava a semplificare la farraginosità della tecnica Thurstone. Propose così uno strumento (anch'esso, come il precedente, con alcune varianti nella sua fase di sperimentazione iniziale) nel quale i punteggi che avrebbero contribuito a definire l'atteggiamento dei rispondenti non erano più demandati ai giudici, bensì affidati agli stessi intervistati. Si trattava cioè di una scala, come quella riportata in Tabella 7.4, composta da una serie di affermazioni alle quali corrispondono risposte a parziale autonomia semantica. La batteria di affermazioni viene preceduta da una domanda del tipo: "quanto sei d'accordo o in disaccordo con ciascuna delle affermazioni che seguono?". Lo scopo è, ancora una volta, quello di intercettare un concetto ad alto livello di generalità. Nell'esempio riportato nella Tabella 7.4 il concetto di riferimento è quello di xenofobia.

**Tabella 7.4** Esempio di scala Likert (adattata da Pavsic e Pitrone, 2003, p. 71)

	<b>Com- pleta- mente in di- sac- cordo</b>	<b>Piut- tosto in di- sac- cordo</b>	<b>Né d'ac- cordo , né in disac- cordo</b>	<b>Piut- tosto d'ac- cordo</b>	<b>Com- pleta- mente d'ac- cordo</b>
Se si accolgono tutti questi immigrati, mancherà sempre più il lavoro per i giovani italiani	○	○	○	○	○
Trovarei naturale che un mio parente stretto sposasse un immigrato di origine africana.	○	○	○	○	○
Tutti questi immigrati hanno fatto aumentare violenza, droga e prostituzione.	○	○	○	○	○

<sup>14</sup> Per il calcolo dell'affidabilità di una scala Thurstone, si veda Perrone (1982, nota 41 a p. 380).

A parità di lavoro, un immigrato ha diritto alla stessa paga di un italiano.	<input type="radio"/>				
Con tutti questi immigrati, le incomprensioni e i conflitti saranno sempre inevitabili.	<input type="radio"/>				
L'Italia deve accogliere tutte le persone costrette a emigrare dalle loro terre di origine per necessità o bisogno.	<input type="radio"/>				
Il matrimonio tra una donna italiana e un immigrato di colore è destinato a finire male.	<input type="radio"/>				
Gli immigrati sono molto più gentili e onesti di tanti di noi italiani.	<input type="radio"/>				
La differenza di salario si giustifica perché gli immigrati di solito rendono meno.	<input type="radio"/>				
L'arrivo di tanti immigrati nel nostro Paese arricchisce la nostra cultura.	<input type="radio"/>				

Le risposte in termini di gradazione di accordo che gli intervistati possono dare alle singole affermazioni, nella formulazione iniziale proposta da Likert, potevano variare da cinque a sette, sebbene tuttora vengano proposte anche scale prive del punto mediano. Generalmente la formulazione del continuum che caratterizza queste risposte varia da “completamente in disaccordo” a “completamente d'accordo”, passando per “lievemente in disaccordo” o “abbastanza in disaccordo”. E già qui cominciano a sorgere alcuni problemi. Poiché la scala – definita dallo stesso autore come una tipica espressione delle *Summated rating scales* (ossia scale additive) – postula l'equidistanza tra i diversi picchetti, il ricercatore è legittimato a pensare che la distanza che esiste tra “completamente d'accordo” e “abbastanza d'accordo” sia la stessa rilevabile tra quest'ultima risposta e “né d'accordo, né in disaccordo” (il picchetto mediano di scala). Ma non è così, poiché – come è stato dimostrato (Amisano e Rinaldi, 1988) – la risposta “abbastanza” viene percepita come maggiormente prossima a *molto* che non a *leggermente* o simili. Anche la scelta di una singola parola, dunque, fa la differenza. Va inoltre tenuto conto che molti ricercatori trattano le scale Likert non come scale ordinali, bensì come se fossero delle quasi-cardinali, il che ha delle implicazioni sulle operazioni in fase di analisi dei dati piuttosto ovvie.

Il punteggio, comunque, viene di solito computato tramite una semplice addizione delle risposte date dagli intervistati: se “completamente in disaccordo” vale 1, “completamente d’accordo” varrà 5 (o viceversa: si tratta di punteggi con valore meramente stipulativo). Più semplice di così... Eppure, per quanto elementare, la procedura di costruzione del punteggio complessivo ha un punto debole: il ricercatore, infatti, sa che soltanto a punteggi tutti estremi di scala corrispondono determinati atteggiamenti (totalmente ostili o a favore del “diverso”, come nell’esempio riportato in Tabella 7.4). Ma i punteggi intermedi non forniscono alcuna indicazione sulle sfumature che caratterizzano le diverse risposte degli intervistati. Vale a dire che se avessimo cinque item e due sequenze di risposte 1 – 2 – 4 – 1 – 2 e 2 – 2 – 2 – 3 – 1, il punteggio dei due soggetti sarebbe identico, ma scaturirebbe da risposte diverse ai singoli item (con la sola eccezione del secondo, dove entrambi rispondono 2).

Un’ulteriore questione rilevante riguarda il problema di come debbano essere scelti gli item da inserire nella scala finale (*cfr.* Capitolo 10). Abbiamo visto che Thurstone e Chave partirono da 130 item sulla Chiesa, per poi selezionare gli undici che servivano, impiegando lo scarto quadratico medio nei casi di *ex-aequo*. Immaginiamo che, nel progettare una scala Likert, l’équipe di ricerca abbia avuto una forma di bulimia creativa, producendo un centinaio d’item. Supponiamo che lo stesso gruppo di studiosi voglia arrivare a collocare soltanto una decina all’interno dello strumento di raccolta dei dati. Come dovrebbe sceglierli? I criteri sono sostanzialmente tre: la coerenza interna della scala; la correlazione tra item; il potere discriminante dei singoli item.

La questione della coerenza interna e quello della correlazione tra item rimandano al problema dell’unidimensionalità della scala. Nelle intenzioni di chi la progetta, la scala dovrebbe essere unidimensionale, ossia fare riferimento a un’unica dimensione concettuale. Se così non fosse, sarebbe opportuno creare tante scale per quante sono le dimensioni, proprio come fecero Adorno e i suoi collaboratori in un celebre studio sulla personalità autoritaria (Adorno, Frenkel-Brunswick, Levinson, Sanford, 1950). Dunque, per accertarsi dell’unidimensionalità della scala, il ricercatore ispeziona innanzitutto le correlazioni tra tutti gli item della scala (*cfr.* Cap. 15). Quelli che mostrano i coefficienti di correlazione più bassi<sup>15</sup> (il che implica un trattamento delle variabili prodotte come quasi-cardinali) salgono immediatamente sul banco dei sospettati. Ma non basta:

---

<sup>15</sup> Gli indici di correlazione variano tra  $-1$  e  $+1$ . I valori “incriminati” sono quelli intorno allo zero.

bisogna anche tenere conto di una misura che suggerisce la coerenza interna della scala. Per farlo, si ricorre a un coefficiente di sintesi, l'Alpha di Cronbach, che si calcola come segue<sup>16</sup>:

$$\alpha = \frac{nr(m)}{1 + r(m)(n - 1)}$$

Nella formula riportata,  $n$  indica il numero degli elementi della scala;  $r(m)$  indica invece la media delle loro correlazioni (cfr. Capitolo 14).

Per capire meglio il procedimento, ci possiamo servire di una scala Likert impiegata in una ricerca mirata a rilevare competenze e sensibilità di un campione di studenti romani delle scuole secondarie di primo grado<sup>17</sup>. Gli item impiegati erano i seguenti: *É indispensabile limitare i consumi*; a) *L'impegno dei cittadini è inutile*; b) *L'inquinamento è il prezzo che paghiamo per il benessere*; c) *La gente si preoccupa troppo dei danni del progresso*; d) *La Natura è governata da forze misteriose*; e) *La questione ambientale non è così drammatica*; f) *La scienza risolverà i problemi ambientali*; g) *Per proteggere l'ambiente servono le leggi*; h) *La cosa più importante è il benessere economico*; i) *É difficile trovare soluzioni per l'inquinamento della Terra*.

Per eseguire il calcolo dell'Alpha di Cronbach, si parte dalla tabella che mette in relazione i punteggi di correlazione degli item. Nella Tabella 7.5 è riportato un estratto della matrice di correlazione inter-item<sup>18</sup>.

**Tabella 7.5** Matrice di correlazione inter-item

	É indispensabile limitare i consumi	L'impegno dei cittadini è inutile	L'inquinamento è il prezzo che paghiamo per il benessere
É indispensabile limitare i consumi	1,000	-,053	,037
L'impegno dei cittadini è inutile	-,053	1,000	,131

<sup>16</sup> Quella riportata è la formula dell'Alpha standardizzato, con valori che oscillano tra 0 e 1. Per un approfondimento, Cfr. Iezzi (2014, p. 145).

<sup>17</sup> Si tratta di una ricerca, a quanto è dato sapere da chi scrive, che l'ex dipartimento RiSMES della Sapienza di Roma realizzò nel 2002 con l'APAT e il CNR.

<sup>18</sup> Riportare la tabella per intero avrebbe significato ottenere una tabella di contingenza di 100 celle (10 x 10).

L'inquinamento è il prezzo che paghiamo per il benessere	,037	,131	1,000
--	------	------	-------

Il secondo passaggio, una volta calcolata la media dei punteggi di correlazione inter-item (pari a 0,27875), è quello di applicare la formula:

$$\alpha = \frac{10 \times 0,04493}{1 + 0,04493 \times 9} = \frac{0,4493}{1,40439} = 0,32$$

**Tabella 7.6** Valori dell'Alpha di Cronbach con esclusione dei singoli item

Item	Alpha di Cronbach se l'item è escluso
È indispensabile limitare i consumi	,373
L'impegno dei cittadini è inutile	,281
L'inquinamento è il prezzo che paghiamo per il benessere	,334
La gente si preoccupa troppo dei danni del progresso	,319
La Natura è governata da forze misteriose	,296
La questione ambientale non è così drammatica	,264
La scienza risolverà i problemi ambientali	,304
Per proteggere l'ambiente servono le leggi	,332
La cosa più importante è il benessere economico	,264
È difficile trovare soluzioni per l'inquinamento della Terra	,324

Al di là del fatto che, complessivamente, il valore dell'Alpha di Cronbach è piuttosto basso (0,32), se si dovesse procedere filologicamente, l'item da escludere inizialmente sarebbe il primo. Cassandolo, infatti, il valore del coefficiente salirebbe a 0,373. Dopodiché si potrebbe estromettere eventualmente il terzo e così via.

L'ultimo criterio del quale tenere conto, nel caso in cui si disponga di un abbondante numero di item e se ne volesse fare una cernita, è il ricorso al potere discriminante dei singoli item. Con questa espressione si intende la capacità che ciascun item possiede di polarizzare le risposte sui perni opposti del continuum che regola l'accordo/disaccordo. Il che è del tutto logico: se le risposte tendono ad assieparsi sui punti centrali della scala, o a concentrarsi su alcune modalità, escludendone altre, l'item con queste

caratteristiche dimostrerebbe di non essere sufficientemente divisivo e, dunque, di non svolgere correttamente la propria funzione.

Il potere discriminante di un item si calcola come differenza dei punteggi medi tra il primo e l'ultimo quartile, ossia il 25% delle risposte collocate sui valori più alti (da "totalmente d'accordo" a scendere) e il 25% di quelle relative ai valori più bassi (da "totalmente in disaccordo" a salire).

Immaginiamo di avere estratto da una scala due item che sono in competizione per trovare un posto nella scala definitiva da somministrare a un campione di intervistati. Supponiamo anche che il pre-test sia stato sottoposto a 36 soggetti. Vediamo come si sono distribuite le risposte dei 9 (= 36/4) che si sono collocati sul versante dell'accordo e dei 9 che hanno espresso il maggiore disaccordo.

**Tabella 7.7** Calcolo del potere discriminante di un item su una scala Likert

	Numero componenti quartile	Totale disaccordo					Totale accordo	Totale ponderato	Media ponderata
		1	2	3	4	5			
Punteggi alti (IV quartile)	9	0	1	2	3	3	<b>35</b>	<b>3,89</b>	
Punteggi bassi (I quartile)	9	1	8	0	0	0	<b>17</b>	<b>1,89</b>	

Il potere discriminante di questo item sarebbe pari a  $3,89 - 1,89 = 2^{19}$ .

Se l'item concorrente avesse valore 1,18, perderebbe automaticamente la competizione proprio perché la sua capacità di polarizzazione si dimostrerebbe inferiore.

Un'altra questione rilevante, implicata nella progettazione delle scale Likert, è quella del *response set*, al quale si è fatto cenno nel paragrafo iniziale di questo capitolo. Per smascherare l'effetto, di norma si consiglia di

<sup>19</sup> I valori 35 e 17 si ottengono dalla somma dei prodotti tra i valori 1, 2, 3, 4 e 5 dei picchetti per il numero di risposte registrate per ciascuno di essi. Dunque, per il IV quartile:  $(0 \times 1) + (1 \times 2) + (2 \times 3) + (3 \times 4) + (3 \times 5) = 0 + 2 + 6 + 12 + 15 = 35$ .

progettare item che non abbiano tutti la stessa polarità semantica<sup>20</sup>. È buona norma anche quella di evitare liste eccessivamente lunghe. Se chi progetta lo strumento di raccolta del dato non intende rinunciare a un folto numero di item, può – al più – spalmarlo in punti diversi del questionario. Chi risponde, infatti, potrebbe – per pigrizia o disinteresse – fornire stancamente sempre la stessa risposta in termini di grado di accordo/disaccordo. Se ciò avviene in un setting di intervista monitorabile dall'intervistatore, a quest'ultimo spetterebbe il compito di richiamare l'attenzione dell'intervistato sulle risposte che sta fornendo, invocandone una maggiore attenzione. Ma se il questionario è autocompilato, la sola strada praticabile è quella dell'estromissione dei casi (o delle sole risposte) che manifestano questa tendenza. Bisogna inoltre fare attenzione al computo dei punteggi complessivi: poiché alcuni item presentano una polarità semantica invertita, bisogna anche invertirne i punteggi. Nell'esempio riportato in Tabella 7.4, l'item "Il matrimonio tra una donna italiana e un immigrato di colore è destinato a finire male" suggerisce il concetto che vuole richiamare in modo negativo, mentre la frase "Gli immigrati sono molto più gentili e onesti di tanti di noi italiani" lo suggerisce in forma positiva. Il punteggio non può essere una semplice somma. Bisognerebbe trasformare i punteggi dell'uno o dell'altro item: 1 diventerebbe 5, 2 diventerebbe 4 e così via. Una strada più breve, in termini di indicazioni da fornire al software di analisi dei dati, è quella di sottrarre da 6 il punteggio degli item a polarità invertita ( $6-1=5$ ;  $6-2=4$  e così via).

Un problema meno evidente, ma non meno spinoso, presentato dalle scale Likert è quello della *curvilinearità*. Con questo termine si intende la possibilità che alcune affermazioni della scala siano costruite in modo tale da implicare la medesima risposta da parte di soggetti che hanno una visione antitetica del reattivo oggetto dell'item. Il termine deriva dal fatto che «questa distorsione richiama lo spostamento che ha subito la collocazione del soggetto lungo il continuum: la sua posizione si curva, assumendo un andamento a ferro di cavallo» (Pitrone e Pavsic 2003, p. 94). Un esempio canonico è la seguente affermazione: «Il voto è il solo modo in cui la gente come me può avere qualche influenza sulle decisioni politiche». Potrebbe rispondere di essere in disaccordo sia chi ritiene che il voto non serva assolutamente a nulla, tantomeno a influenzare le decisioni

---

<sup>20</sup> La polarità semantica indica il verso (se positivo o negativo) al quale punta ciascuna affermazione (item) della scala. Nell'esempio riportato in Tabella 7.4, i primi due item sono evidentemente in polarità semantica invertita: molto difficilmente, se si è d'accordo con la prima affermazione, lo si sarà anche con la seconda.

della politica, sia chi – all’opposto – è talmente ottimista da pensare che quella non sia la sola strada per incidere sulla politica e che, dunque, limitarsi a quella è riduttivo: due concezioni opposte, medesima risposta. Come suggerisce Mauceri (2003, p. 125), «la formulazione di un item curvilineo, di solito, è strettamente connessa alla scelta di un indicatore del concetto [...] che presenta una parte estranea piuttosto estesa ed è proprio questo fatto a renderlo scarsamente discriminante. La distorsione, infatti, risiede nel fatto che il ricercatore interpreta l’item come indicatore di un atteggiamento orientato in una certa direzione (a favore o a sfavore dell’oggetto), quando, invece, in un numero più o meno esteso di casi, gli intervistati, interpretandolo dal loro punto di vista, vi leggono indizi di un atteggiamento orientato nella direzione opposta».

Le ultime due questioni sulle quali il ricercatore dovrebbe porre attenzione e cautela riguardano l’*acquiescenza* e la *reazione all’oggetto*, entrambe smascherabili alla sola condizione che il setting di ricerca implichi la presenza dell’intervistatore. L’acquiescenza esprime la tendenza di un intervistato ad esprimere accordo a prescindere dai contenuti delle affermazioni che gli vengono sottoposte. Si tratta di un atteggiamento che può trovare le sue radici nei «tratti di personalità del soggetto che per timidezza, insicurezza, paura del conflitto, tende a essere sempre d’accordo con qualsiasi affermazione proposta» (Montesperelli, 1998, p. 99).

La reazione all’oggetto riguarda infine la possibilità che l’intervistato, anziché valutare il contenuto dell’affermazione che gli viene sottoposta, si concentri soltanto su alcune parole-chiave che attirano la sua attenzione. Come chiarisce Marradi (2002, p. 36), «una quota cospicua di interrogati non reagisce alle affermazioni, ma ai personaggi, alle azioni, alle situazioni menzionate dalle affermazioni stesse». Il fenomeno, continua Marradi (*ibidem*, p. 37) «è più facilmente individuabile quando la frase di una scala Likert implicitamente o esplicitamente disapprova dei comportamenti, azioni o situazioni che anche l’intervistato disapprova», sicché «il disaccordo riguarda i personaggi o comportamenti oggetto della frase, non la frase stessa» (*ibidem*). È il caso di questa affermazione: «i politici sono interessati solo al voto degli elettori, non ai loro bisogni e alle loro opinioni». Ascoltando le ragioni sottostanti alle risposte, Marradi si accorse che il disaccordo era pilotato dal disprezzo nei confronti della classe politica. Paradossalmente, le risposte avrebbero dunque dovuto far registrare l’accordo, non il contrario.

In breve, per quanto le scale Likert presentino una notevole facilità di costruzione, esse nascondono delle trappole che soltanto l'esperienza del ricercatore e la perizia dell'intervistatore possono tamponare o eludere: il *response set*, l'acquiescenza, la curvilinearità, la presenza di un doppio oggetto all'interno dell'item e la mancata unidimensionalità della scala. Quest'ultimo problema rimanda a delle soluzioni possibili, ma assai complesse e fuori dalle ambizioni di un volume di base per la ricerca sociale: le scale Mokken e Rasch o la scala TLL, per le quali si rimanda a Giampaglia (1990) e Cannavò (2003).

Quanto alla presenza del doppio oggetto, questa circostanza mette l'intervistato nella condizione di non sapere a cosa fare riferimento e produrre una risposta che – nonostante l'eventuale buona volontà – a conti fatti risulta inintelligibile in sede di analisi dei dati. Un esempio canonico sono affermazioni come le seguenti:

*Lo Stato non fa abbastanza per risolvere i problemi di giovani e anziani,  
che non sono nel mercato del lavoro.  
I miei genitori mi hanno fornito un'educazione religiosa.*

Nel primo esempio, chi risponde potrebbe avere una diversa opinione su ciò che lo Stato fa rispetto ai giovani e agli anziani; nel secondo, è possibile che uno dei due genitori sia religioso, ma l'altro no. Come potrebbe rispondere adeguatamente l'intervistato, se non – magari – improvvisando una sorta di media mentale tra i due?

## **5. Le scale autoancoranti e il differenziale semantico**

---

Un meccanismo di computazione dei punteggi assai simile a quello delle scale Likert è rintracciabile in una famiglia di scale che si definiscono autoancoranti. Come si è visto nelle pagine precedenti, l'autoancoraggio si riferisce al fatto che le modalità di risposta della scala sono fissate su due poli opposti di un continuum (per lo più di numeri naturali) che esprimono il massimo e il minimo. Due esempi tipici di scale autoancoranti sono la scala Cantril e il termometro dei sentimenti. La prima è una scala ideata dallo psicologo americano Albert H. Cantril (1946) che può

essere applicata tanto ad affermazioni come quelle di una scala Likert (transitando così da un livello ordinale a uno quasi-cardinale di scala, con le relative implicazioni metodologiche in termini di analisi dei dati), quanto a semplici oggetti o elementi. A titolo di esempio, possiamo immaginare un'indagine sui residenti nei diversi capoluoghi di provincia italiani nella quale si domanda agli intervistati di esprimere un giudizio, su un range che – come si è detto – oscilla tra 1 e 10, rispetto ad alcune caratteristiche della città nel quale risiedono.

La scala si presenterà in questa forma:

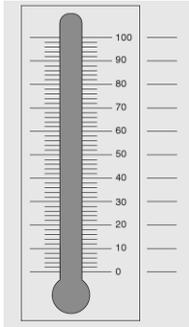
*Riflettendo sulla città nella quale abita, quale valore assegnerebbe a ciascuno di questi aspetti, tenendo conto che il valore 1 indica la minima soddisfazione e 10 la massima soddisfazione nei confronti di essi?*

Gestione della microcriminalità	1	2	3	4	5	6	7	8	9	10
Presenza di aree di verde pubblico	1	2	3	4	5	6	7	8	9	10
Capillarità dei servizi pubblici	1	2	3	4	5	6	7	8	9	10
Efficienza del trasporto pubblico	1	2	3	4	5	6	7	8	9	10
Gestione del decongestionamento del traffico	1	2	3	4	5	6	7	8	9	10
Condizione delle strade	1	2	3	4	5	6	7	8	9	10

Il punteggio complessivo per ogni singolo rispondente si otterrà dalla somma dei punteggi, come nella scala Likert.

La scala Cantril può essere impiegata anche soltanto per ottenere informazioni su singoli soggetti o caratteristiche, per esempio proponendo all'intervistato un elenco di uomini politici rispetto ai quali rilevare il proprio apprezzamento oppure chiedendogli di indicare l'importanza che assumono alcuni tratti della formazione della propria personalità.

Una sorta di variante della scala Cantril è il termometro dei sentimenti. Come suggerisce il nome, si tratta di uno strumento che registra la “temperatura” degli atteggiamenti in una scala graduata che, di norma, ha un centinaio di picchetti. Lo strumento, che di regola viene impiegato in una



**Figura 7.1** Termometro dei sentimenti

forma grafica come quella riportata in Figura 7.1, fa riferimento alla scala di temperatura in gradi Celsius, per cui allo 0 corrisponde la massima freddezza rispetto all’oggetto dell’atteggiamento e a 100 il massimo calore. Per usarlo, il rispondente deve poter disporre dello strumento fisico, di norma, raffigurato su un cartoncino, anche se le risorse del web 2.0 hanno reso possibile un suo travaso – e dunque una sua applicabilità – anche in digitale. Quale che sia il supporto sul quale è fissato il termometro, l’intervistato deve segnare l’intensità che prova rispetto all’oggetto di analisi che gli viene proposto. Se si tratta di affermazioni in batteria, il punteggio complessivo del singolo intervistato sarà conseguito, ancora una volta, per via additiva. Ma il termometro dei sentimenti può essere impiegato anche per registrare, per esempio, il livello di fiducia o sfiducia nei confronti di “oggetti” come istituzioni, singoli uomini politici, partiti, squadre di calcio e via dicendo.

Se il sicuro vantaggio di questo strumento è quello di graduare più analiticamente degli strumenti precedenti l’intensità dell’atteggiamento, esso non è immune dagli stessi difetti che già sono stati indicati: response set, acquiescenza, reazione all’oggetto, curvilinearità. Va inoltre considerato che – alla prova empirica – spesso si verificano due tendenze da parte dei rispondenti: quella a segnare valori che coincidono con le decine (10, 20 ... 80, 90, per esempio) e quella di usare poco o per nulla i valori estremi di scala. Quest’ultimo difetto accomuna il termometro dei sentimenti alla scala Cantril, la quale – ricalcando sostanzialmente i valori ai quali

veniamo socializzati fin dai tempi della scuola – crea un *bias* per cui nel rispondere gli intervistati percepiscono gli estremi come zone *off limits*.

L'ultimo strumento che prenderemo in considerazione è il differenziale semantico, le cui caratteristiche, come ha sottolineato Cannavò (2003, p. 24), «ci impediscono di considerarlo a pieno titolo una scala». Si tratta, infatti, di uno strumento che risponde fin dal principio a una diversa logica rispetto a quelli considerati finora. Esso, infatti, è concepito dal suo ideatore (Osgood, 1952) già in origine, come uno strumento multidimensionale. Ciò sarebbe il meno, visto che – come si è osservato – le scale di atteggiamento considerate in precedenza aspirano all'unidimensionalità, ma non sempre riescono a rilevare i concetti riuscendo a impedire l'ammissione di una cospicua parte estranea (*cfr.* Cap. 3) e, di conseguenza, a esprimere una latente multidimensionalità. La questione più rilevante, in termini procedurali, è un'altra. Mentre tutti gli strumenti precedentemente passati in rassegna approdano a un punteggio complessivo rispetto al singolo rispondente, nel caso del termometro dei sentimenti la traiettoria è diversa. Per capirne il perché, è necessario spiegare cosa sia e come funzioni un differenziale semantico. Partiamo dal nome: esso rimanda all'idea che lo strumento sia formato da una serie di aggettivi antinomici (cinquanta e sempre le stesse, nelle indicazioni dei progettisti) graduate su un continuum che, di regola, è a sette picchetti: da -3 a +3, con lo zero come punto centrale (si veda la Tabella 7.8). Anche in questo caso esistono delle varianti (a cinque picchetti o persino a sei, senza elemento centrale). Stando alle istruzioni per l'uso, il differenziale semantico dovrebbe essere impiegato dagli intervistati come un reattivo cognitivo, al quale si deve rispondere con immediatezza, a sensazione, senza pensarci su troppo. Da ciò emergerebbero tre dimensioni latenti, che sono valutazione, potenza e attività<sup>21</sup>. Le dimensioni latenti sono state ripetutamente osservate empiricamente mediante analisi fattoriale (*cfr.* Cap.17).

---

<sup>21</sup> Per capire cosa si intenda per valutazione, potenza e attività, è necessario fare riferimento alla dimensione semantica dei diversi sottoinsiemi che compongono il differenziale semantico. Coppie di aggettivi come buono/cattivo, bello/brutto, dolce/amaro, utile/inutile, gradevole/sgradevole, pregevole/spregevole, giusto/ingiusto, desiderabile/indesiderabile rimandano all'idea di una valutazione; grande/piccolo, potente/impotente, forte/debole, alto/basso, resistente/fragile, profondo/superficiale richiamano invece una sottostante dimensione di potenza (o di intensità, se si preferisce); infine, le coppie come veloce/lento, vivo/morto, rumoroso/silenzioso, giovane/vecchio, stabile/instabile, preciso/vago, prevedibile/imprevedibile richiamano il concetto ombrello di attività, poiché hanno tutti a che fare con un'idea di azione.

**Tabella 7.8** Differenziale semantico di Osgood

buono	-3	-2	-1	0	+1	+2	+3	cattivo
bello	-3	-2	-1	0	+1	+2	+3	brutto
dolce	-3	-2	-1	0	+1	+2	+3	amaro
utile	-3	-2	-1	0	+1	+2	+3	inutile
gradevole	-3	-2	-1	0	+1	+2	+3	sgradevole
pregevole	-3	-2	-1	0	+1	+2	+3	spregevole
giusto	-3	-2	-1	0	+1	+2	+3	ingiusto
desiderabile	-3	-2	-1	0	+1	+2	+3	indesiderabile
grande	-3	-2	-1	0	+1	+2	+3	piccolo
potente	-3	-2	-1	0	+1	+2	+3	impotente
forte	-3	-2	-1	0	+1	+2	+3	debole
alto	-3	-2	-1	0	+1	+2	+3	basso
resistente	-3	-2	-1	0	+1	+2	+3	fragile
profondo	-3	-2	-1	0	+1	+2	+3	superficiale
veloce	-3	-2	-1	0	+1	+2	+3	lento
vivo	-3	-2	-1	0	+1	+2	+3	morto
rumoroso	-3	-2	-1	0	+1	+2	+3	silenzioso
giovane	-3	-2	-1	0	+1	+2	+3	vecchio
stabile	-3	-2	-1	0	+1	+2	+3	instabile
preciso	-3	-2	-1	0	+1	+2	+3	vago
prevedibile	-3	-2	-1	0	+1	+2	+3	imprevedibile

Nella tabella 7.8 è riportato un ritaglio (per ragioni di spazio) del differenziale semantico (Osgood, Suci, & Tannenbaum, 1957). Si noter  che sulla sinistra sono collocati tutti aggettivi positivi, mentre sulla destra tutti quelli negativi (almeno comparativamente). All'intervistato lo strumento non viene proposto in questo modo: gli uni e gli altri si alternano, per attutire – ancora una volta – la possibilit  che si verifichino casi di *response set*. A risposte raccolte, non necessariamente si perviene a un punteggio del singolo intervistato<sup>22</sup>, come in tutti gli altri casi, bens  a identificare una sorta di trama che sia il singolo rispondente, sia l'insieme degli intervistati fornisce al singolo oggetto di valutazione.

<sup>22</sup> In realt , questa opzione   realizzabile in due modi. Quello pi  semplice   la somma delle risposte date (tenendo sempre conto dell'orientamento positivo o negativo della coppia antinomica) sulle tre dimensioni (valutazione, potenza, attivit ). Non uno, dunque, ma tre distinti punteggi. Il secondo modo per ottenere i punteggi   pi  complesso e richiede il ricorso ai *factor scores* che emergono dalle dimensioni latenti, ma si tratta di una pratica tutt'altro che di uso comune.

**Tabella 7.9** Valutazione dei partiti

Corrotti	○	Onesti	
Buoni		○	Cattivi
Brutti	○	Belli	
Dolci		○	Amari
Precisi	○	Vaghi	
Imprevedibili	○	Prevedibili	

Ne scaturisce qualcosa di assimilabile a quanto riportato nella Tabella 7.9.9, dove i singoli pallini indicano il valore mediano indicato da un immaginario campione di rispondenti alla richiesta valutazione dei partiti politici nel loro complesso.

La tecnica messa a punto da Osgood e dai suoi collaboratori fu impiegata, in origine, nello studio degli atteggiamenti nei confronti dei neri, della Chiesa e della pena di morte. Ma a tutt'oggi se ne trovano applicazioni che travalicano la mera dimensione della survey, come per esempio nell'analisi del contenuto come inchiesta (*cf.* Cap. 8; Di Franco e Nobile, 1995; Nobile, 2012).

Un ultimo esempio di scala autoancorante, nella quale i perni d'appoggio dello strumento sono rappresentati da due affermazioni contrapposte, è la scala a scelta forzata (*Forced Choice*). Si tratta di uno dispositivo che – nonostante fosse stato introdotto da Banfield (1958) già negli anni Cinquanta – ha faticato per un ventennio prima di cominciare a trovare posto nella ricerca sociale, dove tuttora ha un impiego piuttosto ridotto. La scala a scelta forzata ha infatti diversi pregi, giacché riesce a superare alcuni degli ostacoli che abbiamo visto in precedenza: quello dell'unidimensionalità (viaggiando lungo un continuum di affermazioni opposte, è evidente che la scala non possa presentare altre dimensioni); quello delle risposte date a caso: gli intervistati sono indotti a riflettere sui due item opposti; quello dell'acquiescenza: i soggetti devono collocarsi da una o dall'altra parte del continuum, operando una scelta di campo; quello della reazione all'oggetto, che viene comunque presentato in due formulazioni diverse.

Questa scala è concepita con alcune varianti: tipicamente, essa si limita a presentare due affermazioni contrapposte, come nei due esempi che seguono, il primo dei quali è tratto proprio da Banfield (1958, p. 151):

Cosa è meglio?

*Un uomo simpatico e gentile ma  
che non è un amico fidato* vs. *Un amico fidato ma spesso irritabile  
e scortese*

Scegli l'affermazione che più si accorda con il tuo modo di pensare:

*il lavoro degli immigrati aiuta  
l'economia italiana e la sosteni-  
bilità delle pensioni* vs. *con il loro lavoro, gli immigrati tol-  
gono opportunità agli italiani*

Tuttavia, la stessa scala può essere concepita con dei gradienti che ne caratterizzino il continuum, in modo da permettere agli intervistati di posizionarsi su una dimensione più sfumata rispetto a quella dei due poli, come nel caso riportato di seguito:

*Alla fine, chi lavora duro  
ed è capace, ha successo  
nella vita* □ □ □ □ □ *Alla fine, il successo nella  
vita viene determinato dal  
destino e da circostanze for-  
tunate*

Oltre ai pregi già menzionati, le scale a scelta forzata sono particolarmente duttili e indicate per lo studio dei valori, uno dei temi di più difficile trattazione sul piano empirico.

In chiusura, va ricordato che le scale passate in rassegna vanno sottoposte a controlli di validità e attendibilità/affidabilità (*cfr.* Cap. 10). Vale a dire che il ricercatore deve accertarsi che effettivamente gli item individuati siano adeguati al concetto al quale si riferiscono (validità) e che siano altresì capaci di registrare fedelmente lo stato dei singoli soggetti sulla proprietà che si sta rilevando (affidabilità).

Tutto quanto detto fin qui a proposito dello strumento di rilevazione dei dati nella ricerca quantitativa non può essere pensato come una procedura priva di difetti e intralci. Il ricercatore, prima del varo definitivo del questionario, deve accertarsi che lo strumento funzioni in tutte le sue componenti. Come si è detto all'inizio, le scale rilevano concetti ad alto livello di generalità, per cui il rischio di compiere degli errori è particolarmente elevato. Pertanto, per ridurre questo rischio è necessaria una fase di pretest che serve da collaudo allo strumento. Soltanto allora, dopo

essersi accertati che gli item selezionati colgono adeguatamente i concetti ai quali si vogliono riferire, le scale potranno essere impiegate per la rilevazione definitiva.