This is a pre print version of the following article:

Il piano di campionamento dell’indagine sulla alta affidabilità organizzativa / Lalla, M.. - (2020), pp. 80-93.

[Fondazione Rubes Triva] XL Edizioni
Terms of use:
The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

27/04/2024 09:08

(Article begins on next page)

Capitolo 4

Il piano di campionamento dell'indagine sulla alta affidabilità organizzativa

Michele Lalla

5.1. Introduzione

La strategia di campionamento per condurre l'indagine sull'alta affidabilità organizzativa (HRO) nell'indagine in due aziende di servizio pubblico locale, società A e società B, si è basata sull'esperienza di indagini analoghe condotte nel passato (Golzio *et al.* 2014; Golzio e Lalla 2015).

Nel séguito si descrive la procedura di campionamento adottata per l'indagine. Nel paragrafo 2 si illustrano il calcolo della dimensione campionaria e la stratificazione utilizzata. Nel paragrafo 3 si riportano alcune considerazioni sugli esiti della rilevazione. Nel paragrafo 4 si espongono i procedimenti adottati per determinare i fattori di riporto alla popolazione obiettivo. Nel paragrafo 5 si descrivono gli stimatori delle medie e le relative varianze. Nel paragrafo 6 si forniscono alcune caratteristiche degli errori non campionari in generale e in particolare per l'indagine corrente. Nel paragrafo 7 seguono, infine, le conclusioni con alcune osservazioni sul piano e sulle potenzialità dell'indagine.

5.2. Il piano di campionamento

La dimensione del campione è stata determinata fissando l'errore sulla stima della proporzione, P, della modalità di una data variabile qualitativa binaria o resa dicotoma perché non erano note altre informazioni quantitative sulla popolazione rispetto alle variabili oggetto di misura e da stimare (Y). L'espressione utilizzata di solito è la seguente:

$$n_e = \frac{z_{1-\alpha/2}^2 P(1-P)/e^2}{1+\left(\frac{1}{N}\right)\left[\left(z_{1-\alpha/2}^2 P(1-P)/e^2\right)-1\right]} \tag{1}$$

dove N è la dimensione della popolazione obiettivo, e indica l'errore (assoluto) che si commette nella stima della proporzione P nella popolazione, $z_{1-\alpha/2}$ è l'ascissa della curva normale in cui la funzione di ripartizione vale $(1-\alpha/2)$ con α che rappresenta il livello di significatività desiderato per le stime che si ottengono dal campione, e n_e denota la dimensione del campione risultante dalla precisione desiderata o errore (e) massimo atteso delle stime (Cochran 1977). Il valore del livello di significatività α è stato fissato pari al 5% che implica un valore di $z_{1-\alpha/2} = 1{,}96 \cong 2$.

La dimensione del campione deriva da un compromesso tra la precisione desiderata delle stime e le risorse finanziarie e umane disponibili. Sia C l'ammontare delle risorse disponibili, sia C_0 il costo fisso da sostenere per condurre l'indagine, sia c_u il costo unitario di ogni intervista; allora, la dimensione finale quantificata, n_q, sarà data da

$$n_q = \min\left(n_e, \frac{C - C_0}{c_u}\right). \tag{2}$$

Nel calcolo delle dimensioni campionarie tale equazione non è stata utilizzata *ex-ante* perché le Società non avevano posto problemi sulla dimensione per la natura dell'impegno che l'indagine poneva. In base alle risorse disponibili, infatti, e alla presunta propensione alla partecipazione, la dimensione n_q avrebbe dovuto essere dell'ordine delle 250 unità statistiche (ogni unità è un dipendente), ma si concedeva un ragionevole aumento per soddisfare la precisione statistica nell'applicazione della (1).

La scelta dei valori dei parametri si è basata su standard molto comuni: P=0,5; $z_{1-a/2} = 1{,}96$; e=0,05. In altri termini, si è assunto un errore assoluto pari al 5% della proporzione di un fattore binario incognito della popolazione, P, che si intendeva stimare con un coefficiente di confidenza del 95%.

L'uso delle scale può indure l'interesse a stimare la dimensione campionaria sufficiente a rilevare un determinato coefficiente di correlazione della popolazione, ρ; ossia, si individua una numerosità del campione che sia in grado di rilevare un valore piccolo di ρ. Ci si pone, cioè, la domanda: «Quanto deve essere grande il campione affinché si osservi non casualmente un coefficiente di correlazione campionario, r, molto diverso da ρ?» In altri termini, quale dimensione campionaria è richiesta per avere un intervallo di confidenza del 95% che escluda ρ? Qui il coefficiente di correlazione della popolazione, ρ, viene stimato con il coefficiente di correlazione campionario r. Nell'ipotesi nulla, ρ_{H_0}, che il coefficiente di correlazione della popolazione sia $\rho_{H_0} = 0$, la distribuzione campionaria di r è simmetrica e si può usare la distribuzione della statistica t di Student con (n–2) gradi di libertà

$$t_{(n-2);1-\alpha/2} = r \Big/ \sqrt{(1-r^2)/(n-2)}\,. \tag{3}$$

Il simbolo $t_{(n-2);1-\alpha/2}$ è l'ascissa della curva della distribuzione t di Student in cui la funzione di ripartizione vale $(1-\alpha/2)$ con α che rappresenta il livello di significatività desiderato per le stime che si ottengono dal campione o di un potenziale test statistico. Con pochi passaggi algebrici si ottiene la dimensione campionaria desiderata, che segue da tale approccio, indicata con n_e in analogia all'equazione (1):

$$n_e = 2 + t^2_{(n-2);1-\alpha/2}\left[\left(1-r^2\right)\Big/ r^2\right]. \tag{4}$$

L'espressione contiene la statistica $t_{(n-2);1-\alpha/2}$ che dipende a sua volta dalla dimensione campionaria, ma, per dimensioni superiori a 50, si ha $t_{(n-2);1-\alpha/2} \cong 2$ e si può utilizzare tale valore, che corrisponde all'approssimazione del caso precedente. Si può osservare che con la dimensione prevista uguale o superiore all'ordine delle 250 unità occorre osservare un *r* superiore a 0,12. Per i fini dell'indagine questo risultato è soddisfacente perché le correlazioni attese sono, in genere, piú elevate.

5.2.1. *Campionamento stratificato*

L'accertamento della alta affidabilità organizzativa deve considerare il ruolo dei soggetti all'interno della società, perché le posizioni gerarchiche piú elevate hanno piú incisività sugli aspetti organizzativi e dovrebbero avere una maggiore consapevolezza dei rischi e delle necessità operative; pertanto, si è proceduto a stratificare sulla posizione professionale del personale, cosí definita: dirigenti, quadri, capiufficio, impiegati, operai. L'indice di strato, *h*, varierà, quindi, da 1 a *H*=5 e N_{ch} indica la dimensione della popolazione della *c*-esima società nell'*h*-esimo strato.

Nei primi tre strati (dirigenti, quadri, capiufficio) si è programmato di eseguire un censimento; mentre nei rimanenti due strati (impiegati e operai) si è effettuato il campionamento di un numero di unità corrispondente all'applicazione dell'equazione (1). In ogni società, *c*, e strato, *h*, quindi, si è avuta una dimensione quantificata, $n_{q;ch}$, determinata da

$$n_{q;ch} = \lfloor n_{e;ch} + 1 \rfloor \qquad \text{per} \quad c = \mathrm{A}, \mathrm{B}; \quad h = 1, \ldots, H \tag{5}$$

dove $n_{q;ch}$ è, perciò, il numero di dipendenti da selezionare nello strato *h* della società *c*-esima, $n_{e;ch}$ è la dimensione campionaria ottenuta dall'equazione (1) per la *c*-esima società nello strato *h*, e il simbolo $\lfloor \cdot \rfloor$ indica la parte intera. Si è avuto, cosí, un arrotondamento per eccesso in ogni dominio di studio.

I risultati sono riportati nella Tabella 5.1 per le due società A e B. Si noti che, dato il numero degli impiegati della società A, si è suggerito di eseguire il censimento anche nello strato degli impiegati (*h*=4); infatti, lo scarto tra la numerosità dello strato $\left(N_{\mathrm{A}4}\right)$ e la dimensione del campione $\left(n_{q;\mathrm{A}4}\right)$ si è considerato trascurabile.

Tabella 5.1. Numero di unità statistiche della popolazione (N_{ch}), dimensione campionaria quantificata ($n_{q;ch}$) per le società A e B

h	**Posizione**	$N_{\mathrm{A}\,h}$	$n_{q;\,\mathrm{A}\,h}$	$N_{\mathrm{B}\,h}$	$n_{\mathrm{q};\,\mathrm{B}\,h}$
h1	Dirigenti	5	5	10	10
h2	Quadri	4	4	9	9
h3	Capiufficio	13	13		
h4	Impiegati	69	59	227	143
h5	Operai	330	178	1771	316
•	**Totale**	**421**	**259**	**2017**	**478**

5.2.2. *Costruzione della lista campionaria*

Le unità statistiche campionarie sono state selezionate dagli incaricati interni delle due Società A e B, sulla base delle indicazioni fornite dal consulente. La strategia suggerita riguardava l'estrazione delle unità campionarie di ciascun strato dalle liste anagrafiche degli addetti, suddivise per strato, con il metodo del campionamento sistematico circolare; ossia, senza reimmissione e con probabilità uguali. Agli uffici sono state prospettate diverse possibilità, la piú semplice e ordinaria è la seguente: (a) determinare il passo per strato a valore intero, $K_{ch} = \lfloor N_{ch}/n_{ch} \rfloor \in \mathbb{N}$, dove $\mathbb{N}$ è l'insieme dei numeri naturali; (b) individuare il punto di partenza casuale, ρ_{ch}, determinato generando un numero casuale con distribuzione uniforme discreta in $[1, N_{ch}] \subset \mathbb{N}$; (c) selezionare i dipendenti che nella lista anagrafica occupano le posizioni ottenute tramite l'espressione:

$$\rho_{ch} + (j-1)\,K_{ch} - N_{ch} \cdot 1_{[N_{ch}+1,\infty)}\left[\rho_{ch} + (j-1)\,K_{ch}\right] \qquad \text{per} \quad j = 1, \cdots, n_{ch}\,; \tag{6}$$

dove $1_{[\bullet)}[\cdot]$ è la funzione indicatrice che vale 1, se l'argomento appartiene all'insieme specificato nell'indice, 0 altrimenti. L'estrazione inizia, quindi, dal punto di partenza casuale fornito, ρ_{ch}, e prosegue «lungo» la lista, ricominciando all'inizio dopo la fine della lista. Una alternativa era il metodo degli intervalli frazionari (Cicchitelli *et al*. 1997), ma era stato ritenuto piú complesso per essere applicato correttamente dai dipendenti incaricati.

Per sopperire all'eventuale insuccesso degli incaricati nell'ottenere la partecipazione dei soggetti selezionati, era stato suggerito o un sopra-campionamento o, in alternativa, l'estrazione di una lista «suppletiva», che doveva contenere le unità statistiche dette anche «riserve», per brevità, tra le quali selezionare le sostitute di quelle unità che non si riuscivano a intervistare perché si rifiutavano di rispondere. L'entità della lista di riserva poteva essere fissata intorno al 30% della numerosità campionaria, visto che i dipendenti potevano essere indotti a partecipare con strategie assai stringenti.

5.3. Gli esiti della rilevazione campionaria

Il processo di selezione e raccolta dati ha avuto un successo assai simile nei due casi.

La società A ha ottenuto una scarsa cooperazione perché ha incontrato una forte opposizione interna, che ha condotto a una applicazione del processo di estrazione non sempre coerente con il protocollo sopra descritto, sicché c'è stato quasi certamente un effetto di distorsione da autoselezione iniziale e un rifiuto di collaborare. Tali difficoltà sono emerse solo nello strato degli operai, come si può osservare nella Tabella 5.2, perché solo in esso si è effettuato il campionamento. Si noti, poi, che al momento della rilevazione erano cambiate anche le dimensioni della popolazione per strato, come si può verificare confrontando i dati della prima colonna $\left(N_{\mathrm{A}h}\right)$ della Tabella 5.2 con quelli corrispondenti alla prima colonna della Tabella 5.1. La dimensione obiettivo del campione era di 259 unità statistiche, derivante da una precisione del 5% per una percentuale del 50% con un coefficiente di confidenza del 95% per ogni singolo strato campionato. Il mancato obiettivo comporta una diminuzione della precisione desiderata delle stime e, quindi, come si può osservare nella Tabella 5.2, un aumento dell'errore assoluto dal 5% all'8,4%.

La società B, che ha un numero di addetti pari a quasi il quadruplo di quello della società A, ha riscontrato piú impedimenti e difficoltà di realizzazione dell'indagine, perché una protesta quasi organizzata ha inciso sulla distorsione da autoselezione iniziale con un impatto maggiore della società A, come si può riscontrare sempre nella Tabella 5.2. La dimensione obiettivo del campione era di 478 unità statistiche, derivante da una precisione del 5% per una percentuale del 50% con un coefficiente di confidenza del 95% per ogni singolo strato campionato. Il mancato obiettivo comporta una diminuzione della precisione desiderata delle stime e, quindi, come si può osservare nella Tabella 5.2, si ha un aumento dell'errore assoluto dal 5%: al 7,5% per gli impiegati e al 9,9% per gli operai. Da quanto si è capito, nella società B il campione è stato realizzato su tutto il personale; in pratica, su tutti gli impiegati e gli operai, senza distinzione di qualifica. Ciò comporta che il campione ottenuto è di fatto non stratificato, ma assimilabile a un campionamento casuale semplice senza reimmissione. La procedura di realizzazione ha seguíto, quindi, un protocollo di somministrazione casuale dei questionari, che dovevano essere auto-compilati e non richiedevano sforzi eccessivi; pertanto, c'è stato un problema "politico" di opposizione alla rilevazione. Nonostante gli inconvenienti, si è ottenuta una dimensione del campione di 190 unità, che può ritenersi soddisfacente per quanto concerne la precisione statistica relativa alla numerosità, ma un po' insoddisfacente per la distorsione introdotta, giacché si può sempre argomentare, per convenienza, che si siano auto-selezionati piú rispondenti favorevoli/ sfavorevoli ai protocolli di sicurezza.

Tabella 5.2. Numero di unità statistiche della popolazione (N_{ch}), dimensione campionaria (n_{ch}) ottenuta, e errore risultante (e_{ch}) dalla numerosità rilevata per le società A e B

h	**Posizione**	$N_{\mathrm{A}\,h}$	$n_{\mathrm{A}\text{-}h}$	$e_{\mathrm{A};h}$	$N_{\mathrm{B}\,h}$	$n_{\mathrm{B}\,h}$	$e_{\mathrm{B};h}$
h1	Dirigenti	4	4	0	10	10	0
h2	Quadri	4	4	0	9	9	0
h3	Capiufficio	13	13	0			
h4	Impiegati	56	56	0	227	97	0,075
h5	Operai	330	96	0,084	1771	93	0,099
•	**Totale**	**407**	**173**		**2017**	**209**	

5.4. I fattori di riporto alla popolazione obiettivo

In una popolazione $\wp$ di N unità, sia Y il carattere oggetto di stima (per esempio, il punteggio totale della consapevolezza organizzativa) con una distribuzione statistica incognita e valori $(Y_1, Y_2, \cdots, Y_N)$. Si voglia stimare il totale della Y in $\wp$, dato da $Y = \sum_{i=1}^{N} Y_i$, in base al campione osservato $(y_1, y_2, \cdots, y_n)$, dove y_1 indica il valore osservato di Y nell'unità ottenuta dalla prima estrazione, y_2 indica il valore osservato di Y nell'unità ottenuta dalla seconda estrazione, e cosí via fino all'n-esima estrazione. Gli stimatori che si considerano, in genere, sono lineari del tipo

$$\hat{Y} = \sum_{i=1}^{n} w_i \, y_i \tag{7}$$

dove le quantità w_i, dette *pesi*, non dipendono dal numero d'ordine delle osservazioni, ma possono dipendere dal tipo di campionamento adottato e dall'etichetta che individua l'unità statistica selezionata.

In ogni società, $\wp$, oggetto di studio e suddivisa in H strati, il totale della Y di una variabile qualunque assume la seguente espressione:

$$Y_c = \sum_{h=1}^{H} Y_{ch} \, . \tag{8}$$

Nel piano di campionamento probabilistico adottato, per la singola società, si è generato un campione di n unità estratte senza ripetizione (reimmissione), in cui le unità statistiche sono state estratte con probabilità variabili diverse per strato. Siano $(y_1, y_2, \cdots, y_n)$ le osservazioni campionarie ottenute con l'indagine; siano $(\pi_1, \cdots, \pi_H)$ le probabilità di inclusione delle unità statistiche nei vari strati; allora, lo stimatore corretto del totale, $\hat{Y}$, è dato tradizionalmente da

$$\hat{Y}_c = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \frac{y_{chi}}{\pi_{ch}} \tag{9}$$

che è ottenuto dalla combinazione lineare delle osservazioni campionarie negli H=5 strati con pesi per strato pari a $\left(1/\pi_{ch}\right)$, dipendenti dalle etichette delle unità cui si riferiscono le osservazioni. In altri termini, il peso di una unità i, w_i, è il reciproco della probabilità di inclusione, detto *peso base*. Si applicano spesso «correzioni» per sopperire a varie difficoltà dell'indagine, anche se si introduce una non linearità negli stimatori. Tale statistica è detta anche *stimatore per espansione* perché nel campionamento casuale semplice o autoponderante diventa il prodotto della corrispondente grandezza campionaria moltiplicata per l'inverso della frazione di campionamento: $\hat{Y}_c = \left(N_c/n_c\right)\sum_{hi} y_{chi} = N_c\bar{y}_c$. La frazione di campionamento è, in totale, n_c/N_c; quindi, è come se ogni unità del campione rappresentasse N_c/n_c unità della popolazione e, perciò, il fattore N_c/n_c è detto anche *coefficiente di espansione* alla popolazione obiettivo oppure "fattore di espansione all'universo" (Fabbris 1989).

Il peso per strato, calcolato all'origine, $w_{q;\,ch}$, è dato dall'inverso della probabilità di selezione del primo ordine e tutte le unità dello strato h hanno lo stesso peso

$$w_{q;ch} = \frac{1}{\pi_{q;ch}} = \frac{N_{ch}}{n_{q;ch}}. \tag{10}$$

Nel calcolo dei pesi per strato occorre considerare anche la probabilità di ottenere la partecipazione delle singole unità selezionate, $\pi_{p;ch}$; perciò, il peso finale, w_{ch}, è dato dall'equazione (10) trasformata come segue:

$$w_{ch} = \frac{1}{\pi_{ch}} \frac{1}{\pi_{p;ch}} = \frac{N_{ch}}{n_{s;ch}} \frac{n_{s;ch}}{n_{ch}} = \frac{N_{ch}}{n_{ch}} \tag{11}$$

dove $n_{s;ch}$ è il numero di unità selezionate e contattate per invitarle alla partecipazione; n_{ch} denota, come sopra, il numero di unità che partecipano all'indagine e rispondono alle domande del questionario. Il peso finale per strato, ch, corrisponde semplicemente al rapporto tra il numero di unità della popolazione nello strato ch e il numero di unità partecipanti

$$w_{ch} = \frac{N_{ch}}{n_{ch}} = \frac{1}{\pi^*_{p;ch}} \tag{12}$$

dove $1/\pi^*_{p;ch}$ può interpretarsi come una «pseudo-probabilità» di selezione o probabilità di rilevare effettivamente i dati.

Nella Tabella 5.3 sono riportati i pesi, w_{ch}, e la soluzione adottata è la piú semplice per compensare le mancate partecipazioni; altre strategie, piú sofisticate e complesse si trovano in Little e Rubin (1987) e Rubin (1988), ma vanno oltre gli obiettivi di questa indagine. Con la correzione per le mancate risposte gli stimatori diventano, però, non lineari e le loro varianze aumentano (Kish 1990; 1992); inoltre, le correzioni apportate non sono correlate con le variabilità negli strati e tendono a incrementare la varianza; infatti, il peso dei rispondenti aumenta perché «devono rappresentare» anche le unità che rifiutano di partecipare o che sono irreperibili. In altri termini, ogni metodo di stima campionaria assume il principio che le unità incluse nel campione rappresentano anche le altre unità della popolazione che non sono state selezionate o intervistate; ossia, nel caso in oggetto, ogni unità nel campione rappresenta le $(w_{ch}-1)$ unità della popolazione che non hanno partecipato all'indagine.

Tabella 5.3. Numero di unità statistiche della popolazione (N_{ch}), dimensione campionaria (n_{ch}) rilevata, e peso (w_{ch}) per le società A e B

h	**Posizione**	$N_{\mathrm{A}h}$	$n_{\mathrm{A}h}$	$w_{\mathrm{A}h}$	$N_{\mathrm{B}h}$	$n_{\mathrm{B}h}$	$w_{\mathrm{B}h}$
h1	Dirigenti	4	4	1	10	10	1
h2	Quadri	4	4	1	9	9	1
h3	Capiufficio	13	13	1			
h4	Impiegati	56	56	1	227	97	2,3402
h5	Operai	330	96	3,4378	1771	93	19,0430
•	**Totale**	**407**	**173**		**2017**	**209**	

5.5. Varianza delle stime

In un piano di campionamento si riportano, in genere, le espressioni di uno o piú stimatori (media e totale di alcune grandezze) e le relative varianze per documentare l'entità dell'errore di stima. Qui si è applicata una procedura standard e molto documentata nei testi di campionamento, perciò, si riporterà solo qualche sintetica espressione formale. Per esempio, sempre per la caratteristica Y di natura continua, sia $\overline{Y}_c$ la media della c-esima società e $\overline{Y}_{ch}$ la media della c-esima società nell'h-esimo strato, sia W_{ch} il peso dell'h-esimo strato nella c-esima società definito come $W_{ch} = N_{ch}/N_c$, siano $\hat{\overline{y}}_{ch}$ e $\hat{\overline{y}}_c$ gli stimatori della media della c-esima società $\overline{Y}_{ch}$ e $\overline{Y}_c$, rispettivamente, allora l'espressione dello stimatore della media della c-esima società, $\hat{\overline{y}}_c$, è dato da

$$\hat{\overline{y}}_c = \sum_{h=1}^{H} W_h \hat{\overline{y}}_{ch} = \sum_{h=1}^{H} \left(N_{ch}/N_c\right) \hat{\overline{y}}_{ch} . \tag{13}$$

In generale, in un campionamento stratificato, lo stimatore della media della popolazione è non distorto e la varianza dello stimatore della media è data

$$V\left(\hat{\bar{y}}_c\right)=\sum_{h=1}^{H} W_h^2\, V\left(\hat{\bar{y}}_{ch}\right)=\sum_{h=1}^{H} W_h^2\, \frac{S_{ch}^2}{n_{ch}}\, \frac{N_{ch}-n_{ch}}{N_{ch}-1} \tag{14}$$

dove S_{ch}^2 è la varianza della popolazione, ossia la varianza della *c*-esima società nell'*h*-esimo strato, che può essere stimata con la varianza corretta campionaria della *Y* nello strato *ch*, s_{ch}^2, data da $s_{ch}^2=\sum_{i=1}^{n_{ch}}\left(y_{chi}-\bar{y}_{ch}\right)^2 \Big/ \left(n_{ch}-1\right)$, che è la solita formula per il calcolo della varianza campionaria corretta (Kish 1965; Fabbris 1989).

5.6. Errori non campionari

Le indagini dirette all'accertamento delle attitudini e del comportamento dei soggetti possono risultare complicate, specialmente nell'uso della tecnica del questionario auto-compilato, perché manca la presenza catalizzatrice dell'intervistatore (Couper e Groves 1992; Groves *et al*. 2002). Le capacità degli intervistatori sono fondamentali sia per la qualità dei dati raccolti, sia per ottenere la partecipazione delle unità statistiche, anche se si possono utilizzare diversi accorgimenti per migliorare la rilevazione; per esempio, in questi casi, si possono nominare e formare dei responsabili interni che incoraggiano e seguono le unità selezionate nel campione, ma devono essere scelti all'interno della stessa categoria e non necessariamente in una posizione preminente.

Difficoltà e costrizioni varie possono ostacolare il processo di raccolta dei dati: limitazioni temporali, vincoli logistici, e carenza di risorse umane e un po' di disorganizzazione e/o inesperienza. Tali problemi hanno agito nella presente indagine e hanno imposto limiti insormontabili al miglioramento della raccolta dei dati. In ciò emergono, tuttavia, i tratti fisiologici di ogni piano concreto.

Le cause di errore sono molteplici e non sempre controllabili: *errori non campionari*. Si sono indirizzati tutti gli sforzi nel ridurli perché possono diventare anche preponderanti, rispetto agli errori campionari. Non esiste, tuttavia, una teoria completa degli errori non campionari; pertanto, ogni indagine è un caso a sé e presenta un proprio *profilo dell'errore*. In queste due indagini è difficile tracciare un profilo esatto dell'errore non campionario perché gli attori del processo erano fuori del gruppo di ricerca e interno alle due società A e B, sicché non è possibile ricostruire in dettaglio tutte le fasi e le difficoltà emerse in ciascuna di esse.

5.7. Conclusioni

Dai ritardi osservati nella raccolta dei dati e dalle interazioni intercorse con le due società A e B, si evince che gli aspetti piú critici sono derivati dalle difficoltà di realizzazione, tipiche di ogni esperienza sul campo; per esempio, non si sono rilevate tutte le unità statistiche programmate nei diversi strati in cui si era eseguito il campionamento. Per quanto concerne la sostituzione delle mancate partecipazioni, la raccomandazione espressa alle due società era di non ricorrere a tale pratica perché se da un lato migliora la precisione delle stime, dall'altro genera un aumento della distorsione, perché le unità piú disponibili a collaborare potrebbero avere caratteristiche distintive che inficiano o

distorcono le stime dei parametri della popolazione; ma si era stabilita una autonomia di scelta delle società sulla base delle indicazioni fornite dal consulente.

I dirigenti, i quadri, e i capiufficio hanno partecipato sempre, soprattutto per la consapevolezza dell'importanza dell'indagine, ma anche perché si trovavano in una posizione piú delicata e vulnerabile. Si ricorda che i questionari erano anonimi, ma tale indagine coinvolge il rispondente nelle sue mansioni e naturalmente lo insospettisce, anche quando non c'è motivo di sospettare. Ne consegue che il rispondente può sviluppare attitudini ostative verso l'indagine perché contesta i protocolli esistenti e cova pensieri in contrasto con la dirigenza.

L'indagine sull'alta affidabilità organizzativa nelle aziende tende a suscitare una reazione opposta quasi inevitabilmente e, magari, può generare anche la formazione di piccole o grandi coalizioni contro il processo di indagine. Occorrono motivatori e motivazioni che amalgamino gli animi e ispirino sentimenti positivi e propositivi per il benessere individuale e collettivo al fine di superare le diffidenze e le resistenze. La sicurezza è di tutti e tutti sono coinvolti nel bene e nel male. L'opposizione, infatti, c'è stata in entrambe le società tra gli operai. Tra gli impiegati nella società A si è avuta la partecipazione totale, ma si trattava di un numero esiguo di persone, facilmente coinvolgibili, ma anche divisibili e soggetti a strategie piú o meno coercitive. Tra gli impiegati della società B, che erano molto numerosi, si è riscontrata una assenza di partecipazione confrontabile con quella degli operai; ciò mostra quanto sia difficile verificare e misurare la consapevolezza organizzativa tra i dipendenti, come si è già sopra affermato.

Il risultato conseguito può considerarsi comunque soddisfacente per molti aspetti dell'indagine e dei risultati; tuttavia, l'analisi dei dati potrebbe risultare poco precisa quando si devono usare le proporzioni per piccoli sottogruppi, rispetto alle precisioni desiderate *ex ante* e *ex post*.

Bibliografia

Cicchitelli, G., Herzel, A., e Montanari, G. E. (1997). *Il campionamento statistico*, II edizione, il Mulino, Bologna.

Cochran, W. G. (1977). *Sampling Techniques*, John Wiley & Sons, New York.

Golzio, L. E., Lalla, M., e Manni M. G. (2014). Managing the unexpected: A case study, *Rivista Italiana di Economia Demografia e Statistica*, LXVIII (1), 167–174. Atti della 50-esima Riunione Scientifica della Società Italiana di Economia Demografia e Statistica, Convegno «Trasformazioni economiche e sociali agli inizi del terzo millennio: analisi e prospettive», Roma, 29-31 maggio 2013.

Golzio, L. E., e Lalla, M. (2015). Managing the unexpected: A case study on the management practice and perceptions. In Gaëlle Deharo, Allane Madanamoothoo, Sébastien Point (Eds.), *Les risques en entreprise: dialogues entre la gestion et le droit* (pp. 267–284), PAF – Presses Académiques Francophones, Saarbrücken (DE). Actes du congrès tenu à Troyes (FR), 5 et 6 juin 2014.

Couper, M. P., e Groves, R. M. (1992). The Role of Interviewer in Survey Participation, *Survey Methodology*, 18, 263–278.

Fabbris, L. (1989). L'indagine *campionaria – Metodi, disegni e tecniche di campionamento*, NIS: La Nuova Italia Scientifica, Roma.

Groves, R. M., Dillman, D. A., Eltinge, J. L., e Little, R. J. A. (2002). *Survey Nonresponse*, Wiley & Sons, New York.

Kish, L. (1965). *Survey Sampling*, John Wiley & Sons, New York.

Kish, L. (1990). Weighting: why, when, and how, *Proceedings of the Survey Research Methods Section*, American Statistical Association, 121−130.

Kish, L. (1992). Weighting for unequal P_i, *Journal of Official Statistics*, 8(2), 121−130.

Little, R. J. A., e Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, New York.

Rubin, D. B. (1988). *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons, New York.