



RICERCA

APPROCCIO COMPORTAMENTALE AL NUOVO SISTEMA DI VALUTAZIONE DELLE PERFORMANCE ISPRA

Rapporto di ricerca

Luglio 2022

Approccio comportamentale al nuovo sistema di valutazione delle performance ISPR4

Roma: SNA, 2023
ISBN: 978-88-6449-003-8

SNA - Scuola Nazionale dell'Amministrazione
www.sna.gov.it

Servizio Ricerca, innovazione e strategia
ricerca@sna.gov.it

Pubblicazione non in vendita. Nessuna riproduzione, traduzione o adattamento può essere pubblicata senza citare la fonte.



SNA *Presidenza del Consiglio dei Ministri*
Scuola Nazionale dell'Amministrazione

**APPROCCIO COMPORTAMENTALE
AL NUOVO SISTEMA
DI VALUTAZIONE DELLE PERFORMANCE
ISPRA**

La ricerca *Approccio comportamentale al nuovo sistema di valutazione delle performance ISPRA* è stata realizzata nell'ambito delle attività di ricerca 2020-2021 della Scuola Nazionale dell'Amministrazione (SNA), in collaborazione con ISPRA, sulla base della convenzione sottoscritta il 18 settembre 2020.

Gruppo di ricerca

Fabrizio Cafaggi, Coordinatore Unità di Scienze Comportamentali Scuola Nazionale dell'Amministrazione e giudice Consiglio di Stato - Responsabile scientifico

Francesca Marzo, Supporto alla ricerca Applicazione delle scienze comportamentali alle amministrazioni pubbliche e alla regolazione Scuola Nazionale dell'Amministrazione

con la collaborazione di:

Cecilia Colasanti, ISTAT

Giacomo Sillari, Unità di Scienze Comportamentali LUISS Guido Carli e Scuola Nazionale dell'Amministrazione.

Per ISPRA hanno partecipato allo sviluppo della ricerca:

Antonio Cataldo

Federica Dal Brun

Francesco Lazzarini

Maria Antonietta Reatini

Indice

Executive Summary

Introduzione	11
Capitolo 1	13
<i>Approfondimento dei profili comportamentali nei sistemi di valutazione</i>	
1.1 Analisi dei principali lavori presenti in letteratura	13
1.2 Struttura del processo valutativo e complementarità dei metodi di valutazione	14
1.3 Presupposti teorici e premessa metodologica	16
Capitolo 2	19
<i>Analisi della valutazione effettuata nel 2020</i>	
2.1 Analisi della valutazione dei comportamenti agiti nel 2019	19
2.2 Analisi preliminare della scheda di valutazione e della struttura del processo di valutazione	21
2.3 Approfondimento sui pesi dati agli item	27
2.4 Ipotesi sul comportamento dei valutatori	30
2.5 Indagine empirica per validare l'analisi comportamentale	32
Capitolo 3	35
<i>Proposta di interventi comportamentali per la formulazione delle schede di valutazione</i>	
3.1 Alcune valutazioni preliminari sulle modifiche della scheda per il 2021	35
3.2 Contenere i bias creati dalla formulazione degli item	38
3.3 Indurre la differenziazione modificando la scala	40
3.4 Riduzione del bias of mindlessness con l'espressione di pesi	40
3.5 Primi risultati in base alla preliminare elaborazione dei dati raccolti nel 2021	41
Capitolo 4	45
<i>Resoconto del focus group formativo condotto in SNA per ISPR-A</i>	
4.1 Obiettivo del Focus Group: comprendere il ruolo delle aspettative	45
4.2 Svolgimento del focus group: aspettative rivelate e partecipazione alla progettazione	46
4.3 Risultati dei lavori di gruppo e proposte per il processo di valutazione 2021	47
Capitolo 5	49
<i>Proposta di interventi comportamentali sulla strutturazione del processo di valutazione</i>	
5.1 Riduzione del bias di mindlessness con espressione di preferenze su item incrociati	49
5.2 Affiancamento della valutazione tra pari alla valutazione gerarchica	49
5.3 Aumento della consapevolezza con interventi di formazione mirati	50
5.4 Preliminari modifiche al criterio di differenziazione di giudizio	51
Conclusioni	53
Riferimenti bibliografici	55

Executive Summary

La ricerca “Approccio comportamentale al nuovo sistema di valutazione delle performance ISPRA”, dettagliatamente presentato nel report esteso, ha l’obiettivo di individuare alcuni rilevanti profili del sistema di valutazione della performance ISPRA (Istituto Superiore per la Promozione e la Ricerca Ambientale) interpretabili dal punto di vista comportamentale. In particolare, per ogni problematica evidenziata si è cercato di offrire una soluzione basata su una strategia basata sulle scienze comportamentali.

La non completa efficacia degli strumenti tradizionali di valutazione (consistenti in interventi normativi ed incentivi economici) nell’indurre i comportamenti desiderati (in primis focus sugli item di valutazione e differenziazione nelle valutazioni) ha portato, infatti, alla necessità di individuare alcune soluzioni comportamentali da affiancare a tali strumenti e da attuarsi nell’ambito di una strategia di intervento guidata e monitorata. La metodologia impiegata riflette la convinzione che la formazione laboratoriale che affianca e accompagna le amministrazioni nei percorsi di innovazione organizzativa costituisca uno strumento indispensabile. Le attività di ricerca e formazione svolte si inquadrano nel più generale processo di valutazione caratterizzato dall’uso di diversi strumenti, come indicato dalle Linee guida sulla valutazione della performance del Dipartimento della Funzione Pubblica. Queste propongono una pluralità di metodi di valutazione, tra cui quello della valutazione tra pari e della valutazione degli stakeholder, nonché, la *bottom up evaluation*.

Il Sistema di Misurazione e Valutazione della performance, applicato nel 2020 da ISPRA, ha previsto la valutazione gerarchica e quella tra pari: il progetto formativo della SNA ha, quindi, cercato di valutare comparativamente le due valutazioni. Per motivi interni, legati a una profonda resistenza alla valutazione, ai fini della valutazione 2021 ISPRA ha deciso di sospendere la valutazione tra pari e proseguire solo con quella gerarchica. Per tale ragione, in questo rapporto concentreremo l’attenzione sulle valutazioni compiute dai dirigenti.

A partire dalla considerazione che la misurazione e la successiva valutazione della performance organizzativa e individuale è proprio funzionale alla consuntivazione gestionale dei risultati che viene realizzata a livello organizzativo, si è voluto approfondire in che modo la natura stessa degli indicatori e la loro applicazione nell’ambito del processo di valutazione possono accompagnarsi ad errori percettivi e cognitivi, specialmente quando tali processi debbano incorporare aspettative di condotta dei valutati. La presenza di bias, infatti, distorce il processo valutativo e rischia di ridurre l’efficacia: in altri termini, una valutazione caratterizzata da minori distorsioni cognitive è ragionevolmente più trasparente di una che ne ha di più.

Obiettivo dell’impiego delle scienze comportamentali è proprio quello di mitigare gli effetti distorsivi dei bias, aumentando trasparenza ed obiettività. In particolare, nello studio condotto sono stati individuati i seguenti bias, tra loro collegati, che possono concorrere alla distorsione del processo valutativo: bias mindlessness, effetto alone e leniency error, bias di incorniciamento e bias di ancoraggio, e proposti alcuni possibili interventi comportamentali.

1. Bias di mindlessness

Dallo studio dei dati raccolti dal sistema di valutazione di ISPRA è stato rilevato che un comportamento spesso agito da parte di chi valuta è quello di prendere decisioni senza realmente concentrarsi sull’oggetto della valutazione. In altri termini, anche nel processo di valutazione si osserva un atteggiamento diffuso di assunzione di punti di vista e decisioni senza concentrarsi su ciò che si sta facendo o anche, in alcuni

casi, in maniera intenzionalmente negligente. Questo bias può considerarsi il “rumore” di fondo funzionale al verificarsi degli altri errori sistematici di seguito descritti.

Gli interventi comportamentali proposti per aumentare la consapevolezza e ridurre il mindlessness bias sono stati di due tipi:

- a) Chiedere ai valutatori l'assegnazione di pesi all'importanza di ciascun item da valutare, in relazione al tipo di lavoro da svolgere, anche al fine di rendere la valutazione più omogenea se si guarda alle diverse componenti dell'organizzazione. Nello specifico, si è deciso di provare a ridurre la distorsione indotta da un'assegnazione dei pesi preconstituita inserendo un sistema “*agree or explain*” dove si consente al singolo valutatore di indicare item diversi o diversamente definiti rispetto alla scheda, nonché una diversità dei pesi attribuiti ai singoli item. L'idea è, in altri termini, di rendere endogeni alcuni elementi cruciali, finora considerati in modo esogeno, contenendo anche il rischio della sostituibilità degli item di valutazione e inducendo nel valutatore un senso di responsabilità, stimolandolo a contribuire alla definizione dello strumento di rilevazione delle valutazioni.
- b) Poiché passaggio fondamentale nello sviluppo di consapevolezza è il riconoscimento dell'impossibilità di completa coincidenza tra le aspettative del valutatore, quelle dell'organizzazione e quelle del valutato e poiché il sistema di aspettative è cruciale nella definizione del processo di valutazione (oggetto di valutazione è proprio la conformità o difformità tra la performance e le aspettative del valutatore), si è deciso di procedere ad un confronto aperto tra chi redige le schede di valutazione e i valutatori chiamati ad utilizzare tali schede per valutare, tramite focus group formativi sistematici.

2. Effetto alone e leniency error

Nella valutazione delle performance un primo errore (cd effetto alone) ha a che fare con il fatto che, nel giudicare un individuo, il valutatore è influenzato da una sua caratteristica positiva/negativa e tende, di riflesso, a ritenere che egli abbia altre caratteristiche positive/negative che non sono in realtà collegate alla precedente: un valutato riceve un punteggio alto in ogni aspetto considerato, solo perché particolarmente bravo nello svolgere un singolo compito (o viceversa). Il secondo, invece, è un errore di distribuzione (leniency error), che si riscontra per tutti gli item e per tutte le valutazioni raccolte, per il quale esiste una tendenza a dare valutazioni omogenee spostate verso l'alto. Infatti, la rilevazione mostra un tasso di uniformità verso l'alto delle valutazioni compiute dai dirigenti. Nel contesto del nuovo sistema di valutazione tale uniformità appare particolarmente problematica, dal momento che l'obiettivo perseguito dall'amministrazione è tendere verso una maggiore differenziazione delle valutazioni e diversi incentivi sono stati introdotti al tal fine¹.

Nella ricerca, al fine di limitare l'incidenza delle valutazioni viziate sia dall'effetto alone, sia dal leniency error, sono stati imposti i seguenti interventi comportamentali:

- adozione di una scala da 1 a 10 nella scheda di valutazione, che sostituisca la scala da 1 a 4, per spingere verso una desiderata differenziazione dei giudizi e permettere un superamento di entrambi i bias. Una proposta aggiuntiva, che si potrebbe prendere in considerazione nel caso in cui questo accorgimento non dovesse bastare, è quella di introdurre una scala del tutto inusuale e un valore medio su cui soffermarsi per scegliere se valutare più in alto o più in basso di esso (questo potrebbe essere importante per rendere più consapevole l'eventuale scelta di “essere clementi/severi”).
- Approfondimento del confronto tra le aspettative tramite focus group formativi sistematici, che permettano non solo di interpretare correttamente i risultati del processo valutativo di ogni singolo valutato, ma anche di dare spiegazione all'eccessiva uniformazione verso l'alto delle valutazioni per tutti i valutati. Sempre grazie alla condivisione dei diversi punti di vista, infatti, sarà possibile far emergere la complessità e la potenziale conflittualità del sistema delle aspettative.

¹ Come indicato dalle Linee Guida n. 2/2017 e n. 5/2019 del Dipartimento della Funzione Pubblica.

- Affiancamento al metodo di valutazione gerarchica quello della valutazione tra pari, come strumento di *debiasing*, perché la comparazione tra la valutazione gerarchica e quella tra pari consente di chiarire in che modo la presenza o l'assenza di potere gerarchico incide sui sistemi di credenze e sui bias dei valutatori, dirigenti nel primo caso, colleghi nel secondo. Tale comparazione permette di definire le differenze e di determinare il livello di complementarità tra sistemi di valutazione.

3. Attribute framing (bias di incorniciamento) e bias di ancoraggio

Nel contesto di valutazione considerato, le descrizioni delle performance da valutare, la loro formulazione e struttura, possono influenzare ciò su cui i valutatori focalizzano l'attenzione. Se i diversi item proposti (*teammworking*, qualità tecnica, orientamento al risultato, rispetto delle procedure, collaborazione intra/extra funzionale) vengono collegati tra loro fin dalle definizioni, il valutatore è indotto a non mettere in dubbio tali definizioni e a valutare, quindi, in base ad un "frame" prodotto dalla stessa scheda di valutazione: alcune caratteristiche esplicitamente espresse e collegate negli item di valutazione fungono da fulcro della lettura che il valutatore farà del soggetto da valutare. L'errore sistematico consisterà nel valutare in modo condizionato dalle caratteristiche evidenziate nella presentazione stessa degli item di valutazione.

Congiuntamente a questo errore, potrebbe operare anche un bias di ancoraggio, per cui l'individuo si affiderà ad un punto focale (un'ancora, appunto) come riferimento o punto di partenza, per procedere nella valutazione: è possibile immaginare che collegando due item tra loro, così come indotto dalla loro stessa descrizione all'interno delle schede di valutazione, a seconda di quale sarà la risposta data per prima (e quindi in modo casuale), il valutatore tenderà a dare la seconda ancorandola alla prima.

Per limitare i bias di incorniciamento e di ancoraggio quale diretta conseguenza della formulazione degli item, si è proposta una riformulazione delle definizioni degli item all'interno della scheda. In particolare, vengono proposte alcune integrazioni alle definizioni, precisando che la funzione della definizione sia quella di ridurre l'occorrenza di bias derivanti dalla distonia tra il sistema di credenze organizzative e quello dei singoli dirigenti valutatori. Si è proposto quindi di cambiare le definizioni delle aspettative rendendole meno ambigue, mitigando gli aspetti distorsivi.

Nella Figura 1 è riportato uno schema riassuntivo del percorso di ricerca sviluppato e dei risultati evidenziati, sia in termini di caratteristiche dei bias individuati, sia in termini di proposte di intervento:

Figura 1 – Interventi comportamentali sul breve e sul medio periodo

	Focus sugli strumenti utilizzati		Focus sull'intero processo	
Caratteristiche dei bias	Bias inerenti le definizioni nella scheda	Bias relativi la scelta degli item da inserire nella scheda	Bias inerenti la concezione degli item	Bias inerenti l'atteggiamento nei confronti della valutazione
Interventi comportamentali	Modifiche delle schede di valutazione	<ul style="list-style-type: none"> • Esplicitazione dei pesi ex-ante • Focus group 	<ul style="list-style-type: none"> • Espressione di preferenze su item incrociati • Affiancamento valutazione tra pari e gerarchica 	<ul style="list-style-type: none"> • Interventi di formazione mirata • Affiancamento valutazione tra pari e gerarchica

Mentre per quanto riguarda il focus sull'intero processo, si è potuto immaginare forme di intervento delineate ma su cui sono necessari ulteriori approfondimenti, per quanto riguarda il focus sugli strumenti utilizzati, va segnalato che si è già in parte intervenuto modificando la formulazione delle schede da utilizzare nel 2021, col fine di mitigare i bias inerenti le definizioni. Si è anche potuto procedere a un

confronto finalizzato al contrasto dei bias relativi la scelta degli item da inserire nella scheda, tramite l'attivazione del primo focus group formativo.

Nello specifico, il focus group formativo realizzato ha permesso di evidenziare alcuni punti fondamentali di seguito riportati:

- opportunità di non aumentare il numero degli item.
- necessità di avere un'esplicita definizione degli obiettivi personali dei valutatori.
- possibilità di ricondurre gli item ad alcune macro-categorie di valutazione dell'individuo.
- necessità di mantenere distinti item in parte sovrapponibili e disambiguazione di alcuni item.
- centralità del concetto di "*problem solving*" associato all'elemento di innovatività.
- individuazione del concetto di "autonomia".
- importanza di considerare la differenza di peso degli item.
- opportunità di una disambiguazione dell'item "rispetto delle procedure".

Introduzione

La ricerca “approccio comportamentale al nuovo sistema di valutazione delle performance ISPRA” sviluppata dall’Unità di Scienze comportamentali della SNA in collaborazione con ISPRA, nell’ambito della convenzione sottoscritta il 27 ottobre 2020, ha avuto come obiettivo la realizzazione di attività di formazione strutturata e continua riguardante gli aspetti comportamentali concernente i metodi della valutazione delle performance individuali e delle dinamiche socio-cognitive in atto nei processi decisionali implicati.

Il percorso, articolato in diverse fasi, si è sviluppato a partire da un’analisi statistico-comportamentale del nuovo Sistema di Misurazione e Valutazione delle Performance di ISPRA¹, adottato nel 2020, per l’anno 2019, al fine di approfondire alcune problematiche emerse durante la valutazione.

La prima fase è stata dedicata al *desk study* ed è servita ad approfondire i profili comportamentali teorici e ad analizzare le caratteristiche dei bias individuati nei processi di valutazione. parallelamente è stata svolta l’analisi dei dati raccolti durante la sperimentazione del nuovo Sistema di Valutazione della Performance dell’ISPRA (d’ora in poi “Sistema”) con il quale è stata introdotta in via sperimentale la valutazione tra pari. Tale analisi ha affiancato alla normale elaborazione dei risultati della valutazione, un’analisi aggiuntiva finalizzata all’indagine comportamentale dei due diversi metodi di valutazione implementati nel Sistema nella sua fase sperimentale: la valutazione gerarchica e la valutazione tra pari. È stato così possibile studiare punti di forza e di miglioramento del sistema attuale al fine di proporre integrazioni e modifiche per il processo di valutazione dell’anno 2021.

La seconda fase è consistita nella progettazione di un’indagine empirica tramite questionari ed esperimenti di laboratorio sulla base del modello sviluppato nella prima fase, nella loro organizzazione e realizzazione e nell’analisi strutturata dei dati raccolti. Grazie a tale indagine si sono specificati in dettaglio i processi cognitivi, relativi a valutatori e valutati, che accompagnano le diverse modalità di valutazione e le dinamiche dei bias che possono intervenire.

La terza fase, correlata alla seconda, ha permesso la predisposizione, l’organizzazione e la realizzazione di un focus group rivolto ai soggetti coinvolti nel Sistema, e finalizzato alla restituzione dei risultati dello studio. Obiettivo del focus group è stato quello di stimolare una discussione sulle metodologie di valutazione e sui risultati conseguiti, nonché sulle possibili modifiche che possano consentire un miglioramento del processo valutativo e delle conseguenze da esso derivanti per la comunità di lavoro in ISPRA

Infine, da considerarsi come quarta fase trasversale, il progetto è stato utilizzato come strumento di accompagnamento al nuovo processo di valutazione: ha infatti permesso l’integrazione di elementi comportamentali nel procedimento che si è svolto nel 2021, ha indirizzato la riformulazione delle linee guida per la messa a punto del nuovo Sistema da implementare nei prossimi anni e si prefigura come supporto futuro, tramite la progettazione e l’eventuale reiterazione di focus group formativi che abbiano come argomento alcuni elementi individuati come fondamentali durante le fasi precedenti.

Nei capitoli che seguono riportiamo la descrizione del lavoro svolto e dei risultati ottenuti per le suddette fasi. Nello specifico, vengono approfonditi:

- i profili comportamentali nei sistemi di valutazione (Capitolo 1);
- e le domande di ricerca e le ipotesi sviluppate per l’indagine empirica condotta (Capitolo 2);
- una serie di modifiche proposte al fine migliorare le nuove schede utilizzate dal Sistema nel 2021 (per i comportamenti agiti nel 2020) e da utilizzare nel 2022 (per i comportamenti agiti nel 2021) al fine di indurre piccoli cambiamenti comportamentali nel breve termine (Capitolo 3);
- la descrizione e i risultati del focus group condotto con i soggetti che in ISPRA sono coinvolti nel sistema (Capitolo 4);
- alcune proposte comportamentali da implementare nel Sistema, con l’obiettivo di ottenere il cambiamento desiderato nel medio termine (Capitolo 5).

¹ <https://bit.ly/3zuNLn3>

Capitolo 1

APPROFONDIMENTO DEI PROFILI COMPORTAMENTALI NEI SISTEMI DI VALUTAZIONE

Uno dei principi ispiratori per la valutazione delle performance nella Pubblica Amministrazione (PA) è quello dell'affermazione della cultura della valutazione dipendenti/dirigenti (in linea con quanto avviene anche in altri paesi¹, da realizzarsi sulla base della selettività nell'attribuzione degli incentivi economici e di carriera, in modo da premiare i meritevoli. In tale contesto è facile immaginare che tra le conseguenze della misurazione e della valutazione vi sia un sistema di premi e sanzioni (Gabris & Ihrke, 2000). È dunque necessario ragionare sui possibili percorsi di conciliazione tra tale sistema e l'idea generale di valorizzazione, anche tramite l'individuazione e lo sfruttamento di alcuni bias che possono intervenire in queste dinamiche.

In questa prospettiva, un approccio tramite le scienze comportamentali può risultare utile in ciascuna delle fasi del processo di misurazione e valutazione delle performance, per individuare le migliori strategie e guidare il processo decisionale di ciascun partecipante al processo (valutatore e valutato) (So et al., 2015), ma anche al fine di ridurre quanto più possibile gli effetti distorsivi tipici della valutazione condotta da più valutatori, in atto a causa del fatto che, sebbene attenga al rapporto tra il singolo e il valutatore, l'intero processo si inquadra nel più ampio contesto di valutazione di tutti i soggetti che intervengono nel processo. Qui di seguito presentiamo alcuni elementi centrali di questo approccio, a partire da una sintetica presentazione di ciò che possiamo trovare in letteratura sull'argomento.

1.1 Analisi dei principali lavori presenti in letteratura

Al fine di analizzare motivazioni e meccanismi relativi alla efficacia della valutazione, gli studi, sviluppati nel corso degli anni sulla base di intuizioni derivanti da casi studio, si sono concentrati principalmente, da una parte, sulla progettazione dei sistemi di valutazione (indagando quali condizioni siano più favorevoli per la loro reale attuazione oppure le differenze e le somiglianze rispetto a strumenti come la gestione della qualità totale) (Bowman, 1994; Berger et al, 2013; Barbieri et al., 2017), dall'altra, sul tema della relazione tra i sistemi di valutazione e le percezioni dei dipendenti, le forze motivazionali e le caratteristiche demografiche (Marengo & Pasquali, 2010; Taylor & Westover, 2011; Bruno et al., 2017, Coccia, 2019).

In una prospettiva comportamentale si è sostenuto che, fintanto che le organizzazioni vogliono che i loro sistemi di valutazione delle prestazioni raggiungano gli obiettivi prefissati, è di vitale importanza che i valutatori siano formati per svolgere adeguatamente la loro funzione di valutazione e che in tale contesto la formazione comprenda discussioni sui tipici errori di valutazione e strategie per limitarli (Martin & Bartol, 1986). Allo stesso modo, J. S. Bowman (1999) ha osservato che mentre l'uso dei rating presuppone che i valutatori siano ragionevolmente obiettivi e precisi, errori ben noti si verificano nel processo a causa di limitazioni cognitive quali, ad esempio, la limitatezza della razionalità che non permette di avere una visione completa del lavoro svolto, sia estensivamente (il collegamento dei singoli task messi in atto dal valutato) che in profondità (lo sviluppo di un singolo task nel tempo). In aggiunta, analisi empiriche in psicologia applicata (ad esempio, Thorsteinson et al., 2008) e in economia comportamentale (ad esempio, Furnham & Boo, 2011; Kahneman, 2011) suggeriscono che i punteggi delle prestazioni possono essere

¹ Come mostrato dallo studio comparato Barbieri et al. (2017). Per un riferimento a livello europeo, si veda anche la pubblicazione OECD: <https://bit.ly/3J6l6H2>. Si veda anche il documento della Commissione Europea: <https://bit.ly/3BeQF0r>

distorti proprio a causa di limiti cognitivi che possono generare errori quali, ad esempio, il bias di ancoraggio e l'effetto incorniciamento. Ancora più recentemente, infine, si è evidenziato il contributo fondamentale che un approccio comportamentale alla valutazione delle performance può permettere (Andersen & Larsen, 2016; Andersen & Hjortskov, 2016) e si è approfondito l'intervento di possibili bias nella presentazione delle informazioni riguardanti le performance dei manager della pubblica amministrazione (Belardinelli et al, 2018).

Nonostante ciò, da un'analisi della letteratura si rileva una scarsità di studi empirici sui bias cognitivi dei valutatori che possono sistematicamente influenzare la loro valutazione (Belle et al., 2017). Uno degli obiettivi dello studio del Sistema Ispra è proprio quello di contribuire a colmare questa lacuna nella comprensione dell'incidenza di errori sistematici di giudizio e delle modalità in cui tali errori cognitivi influenzano le valutazioni delle performance individuali nelle organizzazioni del settore pubblico. La premessa metodologica è che la presenza di bias distorca il processo valutativo e dunque ne riduca l'efficacia. Dunque, l'obiettivo dell'impiego delle scienze comportamentali è duplice: eliminare le fonti di bias dalla valutazione ove possibile, mitigare gli effetti distorsivi del bias ove la eliminazione non sia possibile. Una valutazione caratterizzata dall'assenza di distorsioni cognitive è più trasparente di una in presenza di distorsioni cognitive.

Dallo studio dei dati raccolti dal sistema, in particolare, si è rilevato che un comportamento spesso agito da parte di chi valuta è quello di prendere decisioni senza realmente concentrarsi sull'oggetto della valutazione (bias di *mindlessness*).

Due profili sono di particolare importanza ed hanno costituito oggetto specifico di indagine in relazione alla definizione delle caratteristiche (item) oggetto della valutazione e della rilevanza attribuita dai valutatori a ciascun item.

L'ipotesi è che, nonostante la percezione del conflitto dimostrata dalla differenza dei pesi dati ai diversi item, entri in gioco un bias chiamato "effetto alone": nel giudicare un individuo, il valutatore è influenzato da una sua caratteristica positiva/negativa e tende di riflesso a ritenere che egli abbia altre caratteristiche positive/negative che non sono in realtà collegate alla precedente (Nisbett & Wilson, 1977). Esistono diversi studi riguardanti l'effetto alone nella letteratura sulla formulazione di giudizi nella valutazione delle performance (ad esempio, è provato che un impiegato riceve un punteggio alto in ogni aspetto considerato, solo perché particolarmente bravo nello svolgere un singolo compito) (Belle et al., 2017). È anche ipotizzabile la possibilità di un errore di distribuzione (problema che si riscontra per tutti gli item e per tutte le valutazioni raccolte), problema noto in letteratura come "leniency error", per il quale esiste una tendenza a dare valutazioni omogenee spostate verso l'alto (Kane et al., 1995; Ng et al, 2011). Tuttavia, la formulazione di giudizi omogenei su item formulati esplicitamente in contrapposizione tra loro è un fenomeno non affrontato in letteratura.

1.2 Struttura del processo valutativo e complementarità dei metodi di valutazione

Approcciando dal punto di vista comportamentale il problema della valutazione delle performance, si può prevedere di lavorare su alcuni aspetti cognitivi complessi quali a) l'aspettativa (la credenza di un individuo che un certo sforzo condurrà al raggiungimento di un determinato obiettivo) (Castelfranchi & Lorini, 2003; Cohen & Ranganath, 2007; Rowe et al., 2008), b) la valenza (il valore che una persona assegna alla ricompensa che consegnerà in seguito ad un suo comportamento e al lavoro svolto) (Ostinelli, 2005; O'Doherty, 2004) e c) la strumentalità (la credenza di una persona che una determinata ricompensa dipenda da una performance elevata) (Taylor & Westover, 2011; Antoni et al., 2017). Tali aspetti, considerati già in quella che viene chiamata "Expectancy Theory" (Vroom, 1964; Van Eerde & Theirry, 1996) possono essere indagati empiricamente per misurare la motivazione sul lavoro analizzando, da una parte, gli obiettivi che le persone vogliono perseguire e, dall'altra, gli sforzi che sono disposti a fare per ricevere premi e/o evitare sanzioni.

Questa analisi tiene conto del fatto che lo strumento principale usato è stato quello della valutazione gerarchica. Nondimeno, le Linee guida della funzione pubblica propongono una pluralità di metodi di

valutazione, tra cui quello della valutazione tra pari e della valutazione degli stakeholder, nonché la *bottom up evaluation*².

Una prima distinzione che può essere fatta tra i diversi metodi considerati riguarda la relazione tra valutato e valutatore. Nello specifico, la valutazione gerarchica è un metodo fondato sulla separazione tra valutatore e valutato, mentre nella valutazione tra pari la coincidenza tra valutatore e valutato è parziale (nell'autovalutazione vi è coincidenza completa): la dissociazione propria della valutazione gerarchica è presente, sebbene manchi la gerarchia. La comparazione tra la valutazione gerarchica e quella tra pari consente di chiarire in che modo la presenza od assenza di potere gerarchico incide sui sistemi di credenze e sui bias dei valutatori, dirigenti nel primo caso, colleghi nel secondo. Tra gli obiettivi dello studio vi era quello di esaminare in che modo la valutazione tra pari possa consentire di identificare e correggere alcuni bias della valutazione gerarchica. Tale comparazione ha consentito di definire le differenze e di determinare il livello di complementarità tra sistemi di valutazione (Cafaggi, 2020).

Occorre, quindi, considerare separatamente i processi decisionali del valutatore nel contesto della valutazione gerarchica ed in quello della valutazione tra pari. In entrambi i casi è stato fondamentale fare riferimento alle teorie motivazionali delle organizzazioni ed utilizzare gli strumenti concettuali legati all'approccio comportamentale per comprendere lo status quo. I sistemi motivazionali del valutatore gerarchico possono differire da quelli impiegati dai pari nel valutare la performance dei propri colleghi.

In questo quadro le diverse aspettative sono state prese in considerazione. Come detto il sistema di valutazione di ISPRA è specificamente basato sulle aspettative. Oggetto di valutazione è la conformità o difformità tra la performance e le aspettative del valutatore. Lo studio ha messo in rilievo, tuttavia, che alle aspettative del valutatore vanno aggiunte quelle dell'organizzazione che valuta i valutatori e quelle dei valutati. Dunque, sebbene la scheda di valutazione menzioni solo le aspettative dei valutatori l'esito del percorso valutativo è determinato dalla combinazione tra i diversi sistemi di aspettative. Giova precisare che, anche nelle ipotesi in cui le schede di valutazione non facciano, come nel caso di ISPRA, esplicito riferimento alle aspettative, queste svolgono un ruolo fondamentale nel processo valutativo e costituiscono pilastro essenziale dell'architettura comportamentale (Cafaggi, 2021).

Per quanto riguarda un'analisi del meccanismo a 360° (si veda per una spiegazione dettagliata dello strumento Mukhopadhyay, 2006 e Chinotti & Festa, 2008), in particolare, le scienze comportamentali possono essere utili per indagare quali caratteristiche la valutazione assume in base a chi è il soggetto che valuta e in che modo l'integrazione tra i diversi metodi di valutazione consente di ridurre la discrezionalità del valutatore e le distorsioni cognitive in cui incorre. Tra i metodi di valutazione si è scelto di indagare la valutazione tra pari ed il rapporto tra questa e quella gerarchica, con particolare riguardo alle differenze esistenti a seconda che vi sia o no un rapporto di gerarchia tra valutatore e valutato e se dal processo di valutazione derivino conseguenze legate all'attribuzione di premi o l'irrogazione di sanzioni.

Una ricerca sui profili comportamentali può aiutare ad indagare non solo in che modo possono distinguersi i metodi di valutazione dei pari e di valutazione gerarchica ma anche in che modo i bias dei soggetti coinvolti nel processo mutano e quali correttivi possono essere apportati al fine di mitigare le distorsioni cognitive nel processo di valutazione.

In particolare, l'esame comparativo tra valutazione gerarchica e valutazione tra pari consentirà di definire in che modo la seconda possa aiutare la soluzione di distorsioni cognitive relative alla valutazione gerarchica con riferimento alla tendenza confermata dai dati della valutazione sul 2019 che vedono una uniformità di valutazione probabilmente legata al *leniency error* (Kane et al., 1995).

² Cfr. d.lgs. n. 74/2017 recante "Modifiche al decreto legislativo 27 ottobre 2009, n. 150, in attuazione dell'articolo 17, comma 1, lettera r), della legge 7 agosto 2015, n. 124 e alle Linee Guida n. 2/2017 emanate dal Dipartimento della Funzione Pubblica. Le nuove Linee Guida n. 5/2019 sono disponibili al seguente link: <https://bit.ly/3bd0gu5>

1.3 Presupposti teorici e premessa metodologica

Alla base dell'approccio comportamentale utilizzato per il nostro contributo, vi è l'idea che la spiegazione più credibile degli atteggiamenti umani è quella fornita da un modello che ripartisce la mente in due sistemi cognitivi, uno impulsivo (Sistema 1) e l'altro riflessivo (Sistema 2) (Kahneman, 2011). Questi due sistemi cognitivi servono per introdurre due diverse specie di homines: il primo è quello postulato dalla teoria economica classica, l'homo oeconomicus, che appartiene in realtà alla mitologia e non all'antropologia perché è modellato su una teoria empirica sostanzialmente falsa; l'altro è l'homo sapiens, ossia l'agente umano così come è in realtà.

L'approccio comportamentale parte, dunque, da questa distinzione per approfondire il modo di ragionare, di decidere e di agire degli esseri umani: sin dai pionieristici studi di Herbert Simon (1979) sono stati sviluppati modelli di ragionamento umano basati su euristiche (scorciatoie del pensiero che rappresentano l'attivazione del Sistema 1 a scapito del Sistema 2) e i conseguenti bias cognitivi, errori sistematici indotti proprio da queste scorciatoie. L'idea di fondo è che le nostre decisioni sono in gran parte inficiate da errori di ragionamento e quindi di giudizio. Questi difetti comportamentali producono euristiche del giudizio rapide e intuitive che si esercitano a detrimento delle nostre capacità deliberative più razionali: in base allo schema duale di Kahneman e Tversky, la nostra capacità decisionale intuitiva e impulsiva, il Sistema 1, prevale sulla facoltà più razionale e riflessiva, il Sistema 2.

A partire da queste considerazioni, la premessa metodologica è che la presenza di bias distorca il processo valutativo e che, quindi, una valutazione caratterizzata dall'assenza di distorsioni cognitive sia più trasparente di una influenzata da distorsioni cognitive; l'obiettivo dell'intervento sullo strumento di rilevazione della valutazione è duplice: eliminare le fonti di bias ove possibile, mitigare gli effetti distorsivi del bias ove la eliminazione non sia possibile.

Sulla base di questi presupposti, della premessa metodologica e delle evidenze empiriche raccolte, si potranno progettare interventi di nudging³ che agiscano non solo da supporto a chi progetta il sistema di valutazione, superando alcuni ostacoli, ma abilitando anche un incremento della performance individuale, incidendo sull'auto-motivazione e sui conseguenti comportamenti da adottare (Larrick, 2004; Soll et al., 2014).

In particolare, lo studio condotto ha portato alla formulazione di alcuni interventi comportamentali sia sul breve sia sul medio periodo, che riguardino piccole modifiche negli strumenti di rilevazione, da una parte, e interventi più importanti sull'intero processo, dall'altra. Uno schema riassuntivo di questa visione di insieme, che includa un'anticipazione delle soluzioni proposte, è rappresentato dalla matrice raffigurata in Figura 1.

³ Nello specifico, strategie di debiasing. Si veda a riguardo il recentissimo Cantarelli et al., 2020 e, per un approccio alla progettazione di interventi alla pubblica amministrazione in generale, Battaglio et al., 2018; per un'introduzione generale al concetto di nudging, si veda Thaler & Sunstein, 2018, per un'analisi critica si veda Viale, 2018.

Figura 1 – Interventi comportamentali sul breve e sul medio periodo

	Focus sugli strumenti utilizzati		Focus sull'intero processo	
Caratteristiche dei bias	Bias inerenti le definizioni nella scheda	Bias relativi la scelta degli item da inserire nella scheda	Bias inerenti la concezione degli item	Bias inerenti l'atteggiamento nei confronti della valutazione
Interventi comportamentali	Modifiche delle schede di valutazione	<ul style="list-style-type: none"> • Esplicitazione dei pesi ex-ante • Focus group 	<ul style="list-style-type: none"> • Espressione di preferenze su item incrociati • Affiancamento valutazione tra pari e gerarchica 	<ul style="list-style-type: none"> • Interventi di formazione mirata • Affiancamento valutazione tra pari e gerarchica

Capitolo 2

ANALISI DELLA VALUTAZIONE EFFETTUATA NEL 2020

Per presentare l'analisi comportamentale effettuata sulle schede di valutazione usate nella fase sperimentale del Sistema di Valutazione ISPRA, prima di tutto, va considerato che errori e bias possono essere attivati sia in base alla scelta degli item da includere nella scheda di valutazione sia per le definizioni che per tali item vengono formulate (come riportato nella Figura 1, conclusiva del capitolo precedente). In entrambi i casi, l'approccio comportamentale può essere di supporto: nel primo caso, individuando i passaggi che possono risultare forieri di bias e formulando definizioni che aiutino un processo di debiasing, nel secondo, permettendo di rendere più trasparente ed obiettivo il processo di valutazione (anche tramite la definizione esplicita dei sistemi di aspettative con cui la valutazione si confronta). In altri termini, poiché i tentativi di occultare le aspettative possono indurre a distorsioni, esplicitare le aspettative e i pesi aiuta ad oggettivizzare il processo valutativo.

Nella prima sezione di questo capitolo procederemo ad un'analisi del testo contenuto nella formulazione dei quesiti e dei diversi gradi di giudizio, preventivamente concettualizzati ed esplicitamente descritti, per avere un quadro preciso dei possibili ambiti di intervento comportamentale specificamente sulla scheda utilizzata in ISPRA per la valutazione 2020 (senza cioè entrare nel merito della scelta degli item da includere).

Per affrontare, invece, le problematiche inerenti la scelta degli item da inserire all'interno della scheda di valutazione è necessario analizzare il confronto tra sistemi di valore dell'organizzazione e quelli dei singoli individui al suo interno. La complessità di tale fenomeno, così come le possibili soluzioni comportamentali che possono essere proposte per le problematiche che possono evidenziarsi, richiede un approfondimento a parte, che rimandiamo in parte alla sezione successiva, relativa i pesi degli item, in parte al capitolo dedicato al focus group e al confronto delle aspettative.

2.1 Analisi della valutazione dei comportamenti agiti nel 2019

A causa dell'emergenza Covid-19, la valutazione relativa all'anno 2019 è stata realizzata in ritardo rispetto ai piani e, per la precisione, nel mese di luglio 2020.

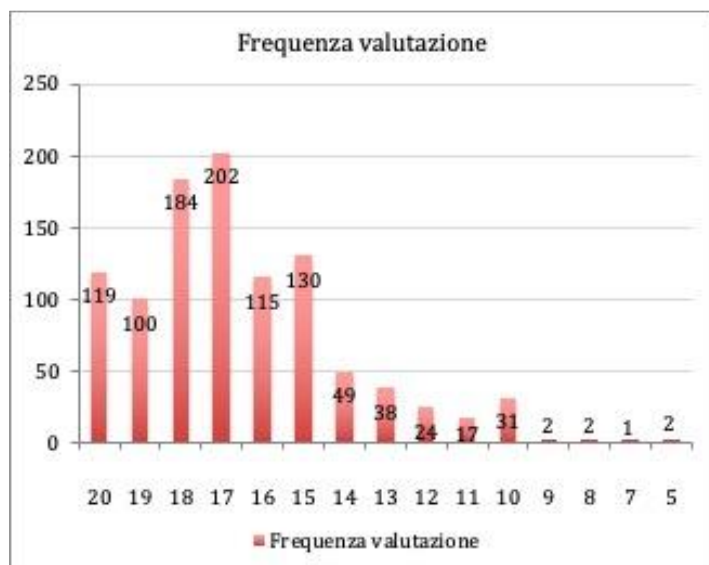
La valutazione è stata effettuata su 1016 dipendenti ISPRA. Un primo elemento significativo per un'analisi del processo è dato dal fatto che una buona parte delle valutazioni effettuate non è stata sottoscritta dai valutati, sebbene una volta concluso il confronto con i valutatori, la metà dei soggetti che non avevano inizialmente sottoscritto hanno finito per accettare la propria valutazione. Va segnalato che una percentuale significativa dei soggetti che non ha accettato, non ha ritenuto di doversi confrontare con i propri valutatori. Entrambi fenomeni danno la misura del segnale di rifiuto dell'intero sistema di valutazione ricevuto dall'organizzazione, che rappresenta una delle ragioni per cui l'indagine è stata richiesta da ISPRA all'Unità di Scienze Comportamentali SNA e avviata.

A questo problema, si è aggiunto il fatto che i dati raccolti dal Sistema confermano una serie di anomalie che rischiano di compromettere l'efficacia stessa del Sistema nel raggiungimento degli obiettivi di miglioramento organizzativo. In particolare, come si evince dalle tabelle riportanti i dati raccolti, qui di seguito mostrate, il problema evidenziato da ISPRA e per cui è stato richiesto un approccio comportamentale riguarda

1. la scarsa differenziazione delle valutazioni;
2. un eccessivo spostamento verso l'alto delle valutazioni;
3. la possibile mancanza di focalizzazione da parte dei valutatori sugli effettivi item da valutare.

Se si compendiano le frequenze con cui le valutazioni sono state compiute (da un massimo di 20 a un minimo di 5), si osservano (Figura 1) i risultati di seguito rappresentati.

Figura 1 – Frequenza con cui le valutazioni sono state effettuate



Fonte: ISPRA

Nello specifico, è possibile osservare che il valore medio è pari a 16,57, il valore più frequente è pari a 17. Per quanto riguarda l'utilizzo della scala di valutazione da parte dei dirigenti valutatori, mantenendo tra 20 e 5 la variabilità, possiamo osservare i valori elaborati in Figura 2.

Figura 2 – Utilizzo della scala di valutazione da parte dei dirigenti valutatori

Valutatore	Valore medio	Deviazione standard
Valutatore#1	15,655	2,663
Valutatore#2	14,974	3,027
Valutatore#3	17,099	1,333
Valutatore#4	16,019	2,461
Valutatore#5	16,794	1,724
Valutatore#6	15,867	1,580
Valutatore#7	15,360	2,343
Valutatore#8	16,836	2,472
Valutatore#9	13,563	3,614
Valutatore#10	17,765	0,903
Valutatore#11	17,026	2,286
Valutatore#12	15,171	3,312
Valutatore#13	16,800	1,473
Valutatore#14	18,441	1,910
Valutatore#15	16,827	2,373
Valutatore#16	16,385	1,121
Valutatore#17	17,100	1,853
Valutatore#18	16,417	2,314
Valutatore#19	17,733	1,911
Valutatore#20	17,368	2,755
Valutatore#21	17,893	2,097
Valutatore#22	17,235	2,658
Valutatore#23	16,737	2,468
Valutatore#24	19,650	0,587
Valutatore#25	18,800	1,229
Totale	16,57	2,54

Fonte: ISPRA

Tali dati acquistano ulteriore consistenza nella segnalazione delle problematiche evidenziate, se si considerano i dati relativi alle valutazioni generali raccolte in aggregato, per tutti i 1016 soggetti valutati, riportati nella seguente Tabella 1.

Tabella 1 – Dati aggregati sulle valutazioni generali (2020)

	X=0	$0 < x \leq 0.25$	$0.25 < x \leq 0,5$	$x > 0,5$
Deviazione standard	249	0	258	509
Media dei giudizi	$x \leq 2$	$2 < x \leq 3$	$3 < x \leq 3.5$	$x > 3.5$
	38	258	317	403

Dai dati riportati qui sopra si possono osservare due cose in particolare: a) che la deviazione standard, indice che mostra la dispersione dei giudizi che i valutatori danno dei singoli valutati, rispetto ai diversi item, permette di ipotizzare un alto numero di valutatori che effettua le valutazioni senza concentrarsi effettivamente sugli item richiesti (deviazione standard uguale a 0); b) le valutazioni sono spostate mediamente verso l'alto, facendo ipotizzare una valutazione finalizzata all'accettazione, piuttosto che al suo efficace utilizzo (che dovrebbe risultare in una maggiore differenziazione tra i soggetti).

2.2 Analisi preliminare della scheda di valutazione e della struttura del processo di valutazione

Nel processo decisionale di valutazione, le stesse descrizioni delle performance di riferimento, così come le schede di valutazione, con la loro formulazione e relativa struttura, possono influenzare sia ciò su cui i valutatori si focalizzano, sia le implicazioni del processo di valutazione. Prima di tutto è necessario dunque evidenziare alcuni aspetti di carattere redazionale che possono indurre bias nel processo decisionale del valutatore.

La scheda di valutazione utilizzata nel 2020 è stata redatta indicando i cinque criteri sotto riportati (Figura 3) a cui è associata una scala da 1 a 4, declinata, nel significato semantico, nelle risposte associate, e nel significato valutativo in 1 = Al di sotto delle aspettative, 2 = In linea con le aspettative, 3 = Sopra le aspettative, 4 = Molto oltre le aspettative.

Dunque, la valutazione non ha neppure la parvenza di oggettività perché viene esplicitamente basata sulle aspettative del valutatore.

Nella scheda di valutazione (Figura 3) non sono indicati i pesi che il valutatore attribuisce a ciascun item. Potrebbe trarsi la conclusione che gli item abbiano tutti lo stesso valore. Ad una indagine svolta dal gruppo di ricerca, cui ha risposto una parte dei valutatori, è però risultato che agli item di valutazione non viene attribuito dai valutatori eguale peso. Oltre alla differenza di peso è emersa anche la diversità di credenze tra i valutatori. Dunque, non solo gli item sono diversamente pesati da ciascun valutatore, ma i pesi relativi sono diversamente determinati dai valutatori. L'importanza relativa degli item che guidano la valutazione è significativamente diversa. Ciò nonostante, le valutazioni risultano omogenee ed appiattite verso l'alto. Ciò dimostra che esiste una contraddizione tra le differenti credenze ed i risultati del processo valutativo, indice evidente della presenza di distorsioni cognitive.

Giova precisare che ISPRA ha deciso di impiegare la stessa scheda di valutazione sia per la valutazione gerarchica sia per quella tra pari (Cfr. Figura 4). Si tornerà su questa scelta e sulle implicazioni che ne derivano in relazione all'obiettivo di valutare differenze e somiglianze tra valutazione gerarchica e paritaria.

Figura 3 – Scheda di valutazione per il personale I-VIII

Unità Operativa	Valutato	Valutatore							
Criteri di valutazione									
					Peso	Valutazione	Punteggio		
Performance organizzativa della struttura dirigenziale di afferenza ex Piano della <i>performance</i> (criterio a.)					85%				
Comportamenti agiti e competenze di ruolo (criterio c.)					15%				
Comportamenti agiti e competenze di ruolo (criterio c.)									
Descrizione competenze e scala di valutazione		1 = Al di sotto delle aspettative 2 = In linea con le aspettative	3 = Sopra le aspettative 4 = Molto oltre le aspettative	<i>Valutazione</i>			Peso		
Teamworking: Capacità di collaborare con i membri del gruppo offrendo il proprio contributo e facendo frutto dell'apporto degli altri per il raggiungimento di uno scopo comune					1	2	3	4	20%
Qualità tecnica: Capacità di esprimere un livello sempre elevato di conoscenza degli aspetti tecnici del lavoro proprio e dei propri collaboratori					1	2	3	4	20%
Orientamento al risultato: Capacità di orientare il proprio sforzo al raggiungimento dell'obiettivo per soddisfare le attese del cliente interno nel rispetto delle regole					1	2	3	4	20%
Rispetto delle procedure: Conoscenza delle procedure interne e capacità di esecuzione delle stesse in modo da non causare ritardi, interruzione o problemi					1	2	3	4	20%
Collaborazione intra ed extra funzionale: Capacità di relazionarsi e costruire rapporti produttivi con i colleghi della propria struttura e delle altre unità operative e con interlocutori esterni all'ente, assicurando un clima di collaborazione.					1	2	3	4	20%
<i>Gli esiti della valutazione sopra riepilogati sono stati comunicati in un "colloquio di feedback" contestualmente alla consegna della presente scheda al valutato.</i>									

Fonte: ISPRA

Figura 4 – Legenda dei comportamenti agiti e delle competenze di ruolo

Teamworking: Capacità di collaborare con i membri del gruppo offrendo il proprio contributo e facendo frutto dell'apporto degli altri per il raggiungimento di uno scopo comune	1= Al di sotto delle aspettative	<i>Si mostra poco collaborativo nei confronti dei membri del gruppo e dimostra di essere produttivo solo quando lavora in solitudine</i>
	2 = In linea con le aspettative	<i>Sa tendenzialmente adattarsi alle dinamiche del gruppo, riuscendo a fornire quasi sempre il proprio contributo</i>
	3 = Sopra le aspettative	<i>E' in grado di collaborare con i membri del gruppo offrendo il proprio contributo e facendo frutto dell'apporto degli altri per il raggiungimento di uno scopo comune</i>
	4 = Molto oltre le aspettative	<i>E' un validissimo membro del gruppo di lavoro, riuscendo a offrire il proprio contributo anche in situazioni difficili mediando e risolvendo possibili situazioni di conflitto tra i membri</i>
Qualità tecnica: Capacità di esprimere un livello sempre elevato di conoscenza degli aspetti tecnici del lavoro proprio e dei propri collaboratori	1= Al di sotto delle aspettative	<i>La non sempre adeguata conoscenza degli aspetti tecnici gli/le impedisce di svolgere alcune delle attività di sua responsabilità</i>
	2 = In linea con le aspettative	<i>Possiede una buona conoscenza degli aspetti tecnici del lavoro che si trova a svolgere nella sua unità</i>
	3 = Sopra le aspettative	<i>Esprime un livello sempre elevato di conoscenza degli aspetti tecnici del lavoro proprio e dei propri collaboratori</i>
	4 = Molto oltre le aspettative	<i>Il livello ottimale di conoscenza degli aspetti tecnici del proprio lavoro gli permette di trovare soluzioni sempre nuove ai problemi raggiungendo risultati superiori alle attese</i>
Orientamento al risultato: Saper attivare e concentrare le proprie energie al fine di raggiungere gli obiettivi in modo concreto e tempestivo	1= Al di sotto delle aspettative	<i>Nella propria attività è orientato esclusivamente al rispetto di regole e procedure anche perdendo di vista l'obiettivo da raggiungere.</i>
	2 = In linea con le aspettative	<i>Riesce tendenzialmente a orientare il proprio sforzo al raggiungimento del risultato per soddisfare le attese dell'utente nel rispetto di regole e procedure.</i>
	3 = Sopra le aspettative	<i>Orienta il proprio sforzo al raggiungimento dell'obiettivo per soddisfare le attese dell'utente nel rispetto di regole e procedure.</i>
	4 = Molto oltre le aspettative	<i>In ogni azione e comportamento ha sempre in mente l'obiettivo e riesce a raggiungere risultati che sono superiori alle attese.</i>
Rispetto delle procedure: Conoscenza delle procedure interne e capacità di esecuzione delle stesse in modo da non causare ritardi, interruzione o problemi	1= Al di sotto delle aspettative	<i>Non conosce tutte le procedure di sua competenza e causa spesso ritardi, interruzioni o problemi nell'applicare le stesse</i>
	2 = In linea con le aspettative	<i>Dimostra una sostanziale conoscenza delle procedure interne, rispettando quasi sempre le scadenze e causando raramente interruzioni o problemi</i>
	3 = Sopra le aspettative	<i>Conosce le procedure interne e fa sì che siano eseguite in modo da non causare ritardi, interruzioni o problemi.</i>
	4 = Molto oltre le aspettative	<i>Possiede un livello ottimale di conoscenza della procedure interne che gli permette di eseguire le stesse in maniera ineccepibile, senza mai causare problemi e ritardi o interruzioni.</i>
Collaborazione intra ed extra funzionale: Capacità di relazionarsi e costruire rapporti produttivi con i colleghi della propria struttura e delle altre unità operative e con interlocutori esterni all'ente, assicurando un clima di collaborazione.	1= Al di sotto delle aspettative	<i>Tende a limitare la collaborazione con i colleghi e con gli interlocutori esterni all'ente, resistendo alla creazione di un clima sereno e collaborativo</i>
	2 = In linea con le aspettative	<i>Collabora in modo abbastanza costruttivo con i colleghi e con interlocutori esterni all'ente, permettendo un clima collaborativ</i>
	3 = Sopra le aspettative	<i>Sa costruire e mantenere rapporti produttivi con tutti gli interlocutori, assicurando un clima di collaborazione.</i>
	4 = Molto oltre le aspettative	<i>Collabora in modo aperto e costruttivo con tutti i suoi interlocutori, favorendo l'instaurazione di un clima di produttiva collaborazione.</i>

Fonte: ISPRA

Procediamo, qui di seguito, ad un esame puntuale della definizione fornita per ciascun item e per i diversi gradi di giudizio. Laddove si evidenzieranno delle problematiche, si ipotizzerà la presenza di diversi specifici bias (per i quali saranno fornite spiegazioni ed esemplificazioni).

Il *teammworking* è definito come la "Capacità di collaborare con i membri del gruppo offrendo il proprio contributo e facendo frutto dell'apporto degli altri per il raggiungimento di uno scopo comune" e la collaborazione intra/extra funzionale come la "Capacità di relazionarsi e costruire rapporti produttivi con i colleghi della propria struttura e delle altre unità operative e con interlocutori esterni all'ente, assicurando un clima di collaborazione".

Tali definizioni presentano aree di sovrapposizione, infatti la "collaborazione intra/extra funzionale" include il concetto di *teammworking*; ciononostante, mentre nella prima l'obiettivo è assicurare un clima di collaborazione, nel *teammworking* l'obiettivo è il raggiungimento di uno scopo comune.

Le aree di sovrapposizione tra i due item possono attivare un bias di rappresentatività, definito come quell'errore in cui siamo indotti da un'euristica che si attiva quando siamo di fronte ad una somiglianza di oggetti o eventi e ad una necessità di dare una probabilità al verificarsi di un evento. Si tratta di una semplificazione euristica dovuta ad un errore nell'elaborazione delle informazioni, per cui non si pensa più rigorosamente, dal punto di vista statistico o logico, ma si formulano giudizi tramite un meccanismo euristico di classificazione delle informazioni basato sull'individuazione, in una particolare istanza, di tratti semantici e prototipici di un concetto: più tratti prototipici identifichiamo in un'accezione, più tendiamo a classificare tale accezione come appartenente ad un concetto generale. Tutto ciò ci permette di essere molto veloci nel classificare oggetti (eventi, persone ecc.), ma può farci facilmente sbagliare (es. grazie a quest'euristica siamo perfettamente in grado di giudicare se, incontrando un gatto, abbiamo di fronte un gatto o un cane, ma allo stesso tempo rischiamo di sbagliare se abbiamo di fronte un animale volatile che allatta, come può essere un pipistrello¹). In altri termini, la somiglianza di oggetti o eventi confonde il pensiero degli individui riguardo alla probabilità di un risultato, inducendoli nell'errore di credere che due cose o eventi simili siano più strettamente correlati di quanto non siano in realtà. Nel caso in esame, se si guarda la scala delle risposte da 1 a 4, sia nel *teammworking* sia nella collaborazione intra/extra funzionale le definizioni sottendono che più si è collaborativi, maggiori sono le probabilità di raggiungere uno scopo comune. Già in due gradi di giudizio di team working (il primo e il terzo, nello specifico) c'è un riferimento alla collaboratività. Ciò può indurre una connessione tra questo item e quello sulla collaborazione intra-extra funzionale, in cui, ovviamente, tutti i gradi di giudizio fanno riferimento alla collaboratività. Il valutatore che utilizza la scheda si può trovare a decidere come valutare il valutato su quest'ultimo item in base a quanto espresso per l'item precedente. Ma giudicare un soggetto sulla sua capacità di collaborare in base a come raggiunge obiettivi in gruppo non è sempre corretto, perché esclude tutte quelle forme di collaborazione (in particolare quelle extra-funzionali) in cui non ci sono necessariamente obiettivi comuni ovvero l'ipotesi in cui la competizione tra appartenenti al gruppo possa consentire il raggiungimento dello scopo comune.

Nel nostro caso particolare, il concetto di collaboratività, in base al tratto semantico prototipico di "raggiungimento di scopi comuni", induce a credere che se si raggiunge uno scopo comune, i soggetti sono altamente cooperativi.

Il problema è che, partendo da tale presupposto, il valutatore, posto in questo evidente collegamento tra item, è indotto a, prima di tutto, a non mettere in dubbio questa relazione concettuale, cadendo in un *attribute framing*, in secondo luogo, a cadere in un potenziale bias di ancoraggio.

Con il primo bias si fa particolare riferimento ad una tipologia di framing effect, tipico errore sistematico che può attivarsi in fase di "editing" del problema (per cui problemi uguali, presentati con valenze differenti, inducono a scelte differenti) (Tversky & Kahneman, 1981). Nel caso dell'*attribute framing*, invece, non la valenza (positiva o negativa) con cui viene presentato il problema, ma alcune caratteristiche esplicitamente espresse di un oggetto o di un evento fungono da fulcro della manipolazione (Levin et al., 1998). L'errore fa sì che se qualcosa viene presentato in un determinato modo, la decisione che sarà presa sarà legata alle caratteristiche evidenziate nella presentazione stessa.

¹ Per un quadro generale dell'euristica in questione per la quale si è indotti a violare regole probabilistiche a favore di opzioni più rappresentative e più disponibili si veda anche il riferimento base Tversky & Kahneman (1972).

Il bias di ancoraggio, invece, si verifica quando un individuo si affida alla primissima informazione appresa per prendere una decisione. Spesso si tratta proprio dell'aggancio ad un punto focale (un'ancora, appunto) come riferimento o punto di partenza, dal quale non ci si distanzia facilmente, finendo con l'aver un grave impatto sulla decisione che l'individuo prenderà (per i meccanismi che sottendono questo comportamento si veda Jacowitz & Kahneman, 1995 e Chapman & Johnson, 2000).

Sebbene l'interazione tra questi due tipi di bias non sia stata specificamente studiata in letteratura e sarebbe interessante capirne le dinamiche anche in contesti più semplici, nel caso in esame, è possibile immaginare che collegando i due item tra loro, così come indotto dal framing, a seconda di quale sarà la risposta data per prima (e quindi in modo casuale), il valutatore tenderà a dare la seconda ancorandola alla prima.

L'obiettivo del lavoro è stato quello di disancorare team working e collaborazione in modo da consentire nella valutazione di dissociare la mancata attitudine alla collaborazione dalla possibile motivazione al raggiungimento dello scopo comune

Per quanto riguarda l'orientamento al risultato viene utilizzata una scala eterogenea, perché introduce in una (e una sola) delle definizioni un nuovo soggetto che è l'utente e un nuovo elemento che è la sua soddisfazione. In particolare, la scala da 1 a 4 del item "orientamento al risultato" fa riferimento grado di giudizio 1 al rispetto di regole e procedure "Nella propria attività è orientato esclusivamente al rispetto di regole e procedure anche perdendo di vista l'obiettivo da raggiungere".

Apparirebbe quindi che, da chi ha redatto la scheda di valutazione, il rispetto delle regole e procedure sia visto come un possibile vincolo per il raggiungimento dell'obiettivo, cadendo così in un bias di correlazione illusoria, in base al quale si inferisce un'associazione tra due variabili (eventi, azioni, idee, ecc.) quando possono non essere effettivamente associate.

Procedendo con questa correlazione che pone implicitamente in contrasto orientamento al risultato e rispetto delle procedure, il valutatore può incorrere in un bias di conferma (Munro & Stansbury, 2009). Quest'ultimo, che consiste nel ricercare, selezionare e interpretare informazioni in modo da porre maggiore attenzione, e quindi attribuire maggiore credibilità a quelle che confermano le proprie convinzioni (e viceversa, ignorare o sminuire informazioni che le contraddicono), può infatti far sì che il valutatore richiami alla memoria solo situazioni e comportamenti del valutato che confermino la correlazione "suggerita" dalla formulazione del quesito.

Dunque, è possibile che il bias di chi ha redatto la scheda, correlando orientamento al risultato e mancato rispetto delle regole, sia stato recepito da chi ha valutato differenziando tra chi è orientato al risultato rispettando le regole e chi è orientato al risultato indipendentemente dal rispetto delle regole. Ma è anche possibile che il valutatore abbia cercato nella propria valutazione, resosi consapevole della associazione implicitamente disegnata da chi ha predisposto le schede di valutazione, di disancorare il raggiungimento del risultato dal rispetto delle procedure ponendo in essere un processo di debiasing². Sempre per l'item orientamento al risultato, nel definire il secondo grado di giudizio, invece, si introduce la soddisfazione delle aspettative dell'utente, dando luogo ad una scala estremamente disomogenea.

La disomogeneità della scala di risposta, creando confusione, può indurre il valutatore a rispondere in modo soggettivo ad un *mindlessness bias* (Luttrell et al., 2014): il valutatore potrebbe ritrovarsi a minimizzare il proprio apporto al sistema di valutazione e operare, perciò, in modo intuitivo e automatico (sistema 1), oppure prendendo decisioni motivate da fattori endogeni (ad esempio fattori emotivi, quali astio personale, invidia ecc.) e valutare, così, in modo distorto.

E' possibile che, in una simile condizione di ambiguità, il valutatore sia soggetto ad un bias di disponibilità, un errore dovuto ad un'euristica in base alla quale le persone esprimono giudizi sulla probabilità di un evento in base alla facilità con cui viene in mente un esempio, un'istanza o un caso (ad

² Il rapporto tra i convincimenti di chi redige la scheda e quelli da chi la usa per valutare è molto complesso. Il valutatore cerca di comprendere quali siano le credenze ed i valori dell'organizzazione. Il problema è capire come questo avviene. Nella discussione si sono ipotizzate tecniche con le quali il valutatore possa chiarire, specificare o anche dissentire dalla formulazione rendendo così esplicito il debiasing. Il focus group ha svolto con successo questa funzione collettivamente stimolando i valutatori ad esprimere i dubbi e le perplessità sulle formulazioni adottate nelle schede.

esempio, in ambito medico, i pazienti possono giudicare la qualità di un intervento sulla base di informazioni che sono apparse di recente nelle notizie, ignorando altri fatti rilevanti) (Tversky & Kahneman, 1973; 1974). Nel nostro caso, a causa dell'ambiguità e della conseguente astrazione della valutazione (l'ambiguità, abbiamo detto, induce ad attivare il sistema 1 e a non fare più riferimento specifico ai singoli item), tale bias potrebbe dar luogo ad una valutazione che non consideri il comportamento del valutato in relazione all'item in esame nel corso dell'intero anno di riferimento della valutazione in corso, ma che faccia riferimento a qualcosa accaduto solo recentemente. Se, ad esempio, un soggetto che deve essere valutato è stato molto orientato al risultato durante tutto l'anno, meritando una valutazione molto alta per questo item, ma recentemente ha avuto problemi a soddisfare i bisogni dell'utenza, il valutatore potrebbe "agganciarsi" a questa "disponibilità" nella sua memoria, dando una valutazione più bassa e pari al secondo grado, solo perché il rapporto con l'utenza è citata in quest'ultimo grado di giudizio. Un episodio negativo recente può influenzare la valutazione in maniera più che proporzionale alla sua incidenza sulla condotta complessiva. L'idea è che il rischio di tale bias possa essere accentuato dalla necessità di cercare occorrenze concrete di comportamenti osservati, nel tentativo da parte del valutatore di uscire dall'ambiguità in cui si trova a causa della formulazione estremamente disomogenea dei gradi di valutazione.

Il contributo all'innovazione e al cambiamento è definito come "Partecipare ad iniziative prese dai propri colleghi e superiori, individuando le direttrici di evoluzione del sapere e creando le condizioni per la relativa acquisizione". Sembra quindi non essere preso adeguatamente in considerazione un approccio proattivo e di responsabilità individuale, tradizionalmente associato al concetto di innovazione, ma si pone l'attenzione solo a cogliere le proposte innovative da parte di altri.

Da tale formulazione sarebbe possibile inferire dunque che l'ambiente presenti una difficoltà generale di ascolto, visto che questa caratteristica viene valutata in modo così peculiare, al punto di influire sulla redazione della scheda di valutazione. Le risposte da 1 a 4 comunque sono coerenti con la definizione data.

C'è un'ambiguità del termine *iniziativa* usato sia in questa definizione, sia come item successivo, che potrebbe acuire, ancora una volta, un attribute framing e il conseguente bias di ancoraggio nel valutatore.

Nell'item *iniziativa*, i concetti di "cogliere opportunità" e "assumersi dei rischi" sono legati appunto alla capacità di iniziativa: nel punto 1, il valutato appare come incapace di cogliere l'opportunità, nel punto 2, invece, è in grado di assumere i rischi; nel punto 3 è esplicitato il concetto di opportunità e di rischi, nel punto 4 viene amplificato il punto 3. Assumere rischi significa assumersi responsabilità. Quando ci si assume un rischio, si valuta più o meno consapevolmente la probabilità di accadimento di una minaccia e l'impatto che ha.

Nella formulazione delle risposte a questo item, il concetto di responsabilità non è esplicito, mentre il concetto di rischio emerge solo in due delle quattro risposte. Di nuovo, come nel caso dell'orientamento al risultato, siamo in presenza di ambiguità. In particolare, in un contesto in cui dobbiamo assumere che "innovazione" significhi semplicemente "ascoltare gli altri", nella risposta 4 viene richiamato il concetto di "innovazione". Non appare chiaro se "innovazione" debba essere intesa come essa viene definita in questo contesto o come invece essa viene intesa più comunemente.

Da segnalare che, in casi come questo, è molto forte il rischio che il valutatore venga influenzato dall'attribute framing: è possibile che una volta letta la definizione di "innovazione" proposta, non riesca più a valutare la capacità di innovazione se il valutato la mostra in modi diversi da quelli sottesi nella descrizione.

Per come sono definiti sia l'innovazione sia la relativa scala di valutazione, si sottende più volte la gradazione con cui si resiste/accoglie il cambiamento, quasi a testimoniare, in chi ha redatto la scheda, un'aspettativa di bias di status quo. Questo bias, che può essere definito come "resistenza al cambiamento", è un errore causato dall'insistenza di un'euristica che si attiva e agisce nella nostra mente in modo molto consistente, poiché si tratta di un meccanismo di "risparmio energetico" evolutivamente sviluppato: qualsiasi spostamento dalla situazione attuale è considerato meno preferibile del non spostarsi affatto (Samuelson & Zeckhauser, 1988, Kahneman et al., 1991). Riportando tale concetto al caso in esame, dalla definizione del concetto di innovazione e dalle descrizioni delle

attitudini di resistere e/o accogliere il cambiamento, sembra che la resistenza al cambiamento venga considerata una situazione di default.

2.3 Approfondimento sui pesi dati agli item

Nella scheda di valutazione non vengono indicati pesi diversi attribuiti agli item. Deve ritenersi che per l'organizzazione essi pesino in modo eguale. Tuttavia, ove invece ai valutatori fosse data la possibilità di attribuire esplicitamente pesi diversi agli item ne deriverebbe una differenziazione del processo di valutazione. L'attribuzione esplicita di pesi da parte dei valutatori potrebbe consentire una differenziazione che consideri anche il sistema di credenze e valori di ciascun dirigente e la sua omogeneità o disomogeneità con quello degli altri valutatori e dell'organizzazione.

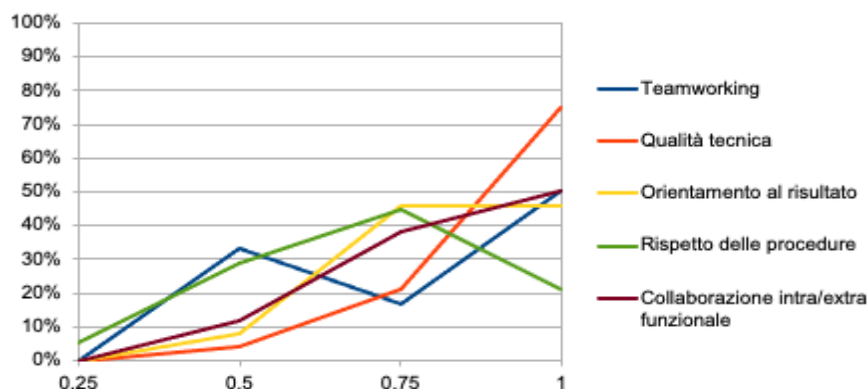
Al fine di comprendere il peso attribuito dai valutatori ai profili oggetto di valutazione (Tabella 2), è stato chiesto ai valutatori di assegnare un peso all'importanza dei diversi item che compongono la scheda di valutazione in relazione al tipo di lavoro da svolgere e in modo astratto rispetto al valutato (per il momento si è deciso di utilizzare questa scala di pesi nella valutazione gerarchica).

Tabella 2 - Pesi espressi dai dirigenti Ispra tramite somministrazioni ex-post di schede dedicate (2020)

Item	Peso assegnato (0,25 è meno importante e 1 più importante)							
	0,25		0,5		0,75		1	
<i>Teamworking</i>			33%		17%		50%	
Qualità tecnica	75%		4%		21%		1	
Orientamento al risultato	46%		8%		46%		1	
Rispetto delle procedure	21%	5%	29%		45%		1	
Collaborazione intra/extra funzionale	50%		12%		38%		1	

Nella tabella sopra riportata è possibile osservare alcune regolarità. In particolare, osserviamo che le preferenze espresse riguardo il *teamworking* sono piuttosto distribuite, con due massimi relativi a 0.5 e 1; il 50% dà importanza 1 ed il restante 50% è diviso tra 0.5 (33%) e 0.75(17%); il dato relativo la qualità tecnica, invece, è molto concentrato sul valore 1 (75%), mentre quasi la totalità della parte restante del campione dà una valutazione di 0.75 (21%) e solo il 4% dà una valutazione di 0.5; il dato riguardante l'orientamento al risultato è piuttosto concentrato sui valori 0.75 e 1 (entrambi al 46%) L'8% dà una valutazione di 0.5; il dato riguardante il rispetto delle procedure è, al contrario, molto distribuito, con un massimo a 0.75 (45%), ma con valori significativi per tutti gli altri punteggi (è l'unica delle voci alla quale è stato dato anche il punteggio di 0.25, anche se solo dal 5% del campione); infine, il dato relativo la collaborazione intra/extra funzionale è con andamento monotono crescente, come per la qualità tecnica, ma più distribuito di quest'ultimo (il massimo è a 1 (50%). Il restante 50% è in buona parte per 0.75 (38%). Il 12% dà una valutazione di 0.5. Per una rappresentazione grafica di queste caratteristiche delle distribuzioni, si veda la figura 5.

Figura 5 - Distribuzione dei pesi espressi dai dirigenti Ispra tramite somministrazioni ex-post di schede dedicate (2020)

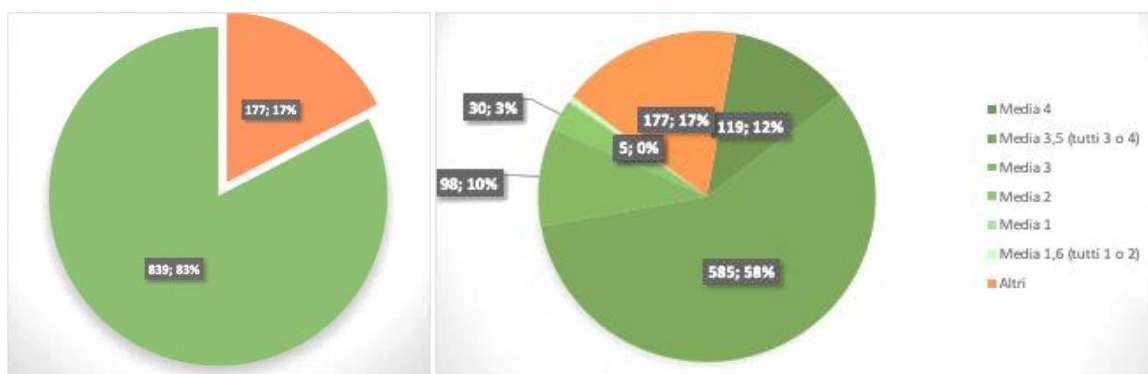


Fonte: elaborazione su dati ISPRA

In generale, ad eccezione della “qualità tecnica”, i punteggi assegnati dai valutatori sono piuttosto distribuiti, anche se il massimo della distribuzione è 1 per tutte le voci, tranne che per il “rispetto delle procedure”. Inoltre, nel caso di Teamworking, Qualità tecnica e Collaborazione intra/extra funzionale, la metà del campione o addirittura più della metà assegna il punteggio 1.

Da questi dati, due profili emergono in modo particolarmente evidente: 1) il peso associato dai valutatori alla competenza (qualità tecnica) è molto elevato rispetto a tutti gli altri, 2) l’orientamento al risultato è considerato un item più importante del rispetto delle procedure. Questa differenza di pesi espressi dai valutatori porta ad ipotizzare che l’ambiguità interpretativa degli item li abbia indotti a percepire l’esistenza di un conflitto tra i due aspetti. A partire dalla formulazione della domanda e dai pesi associati ai due item da parte dei valutatori, dunque, è ragionevole aspettarsi che chi è valutato positivamente nel rispetto delle procedure non lo sia nell’orientamento al risultato. Ciò che invece emerge dai dati raccolti in Ispra, come mostrato in Figura 6, è che l’83% dei valutati è stato giudicato positivamente (valori 3,4), ovvero negativamente, (valori 1,2) in entrambi gli item, in modo quindi uniforme.

Figura 6 - Risultati concordanza dati raccolti dal sistema di valutazione implementato in Ispra nel 2020

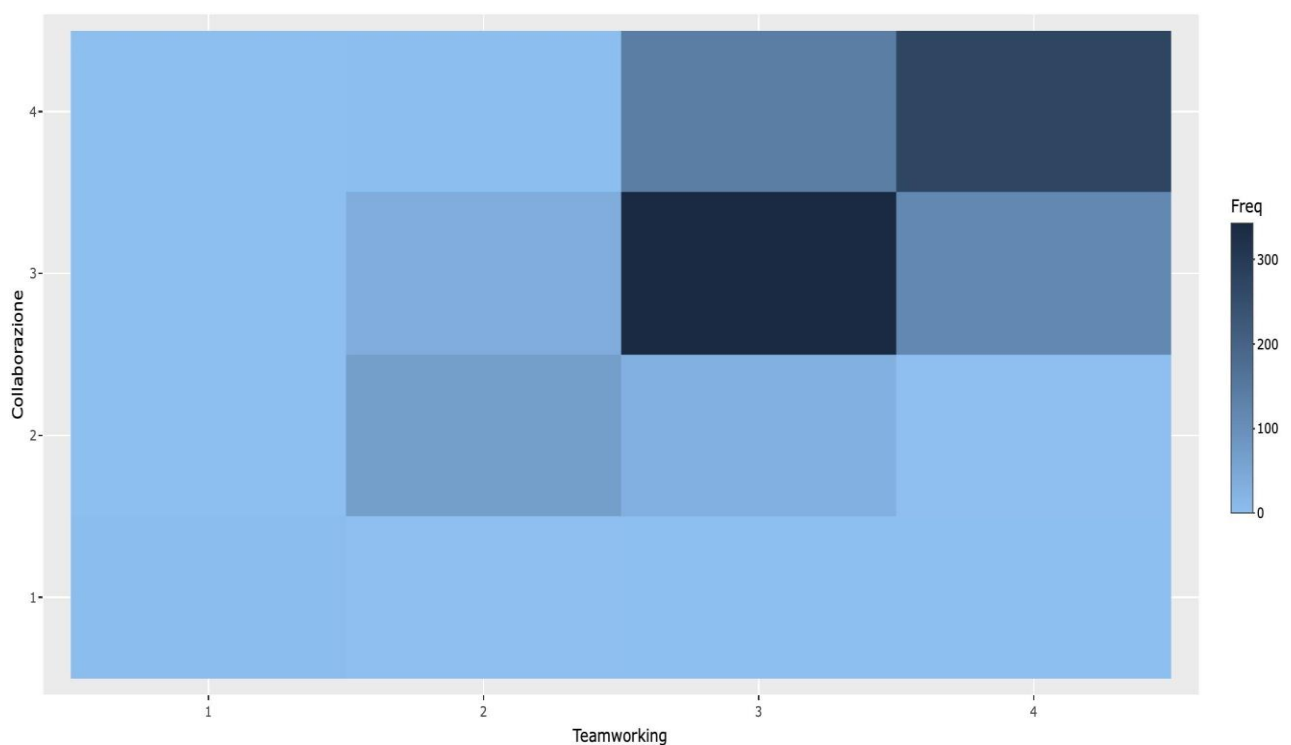


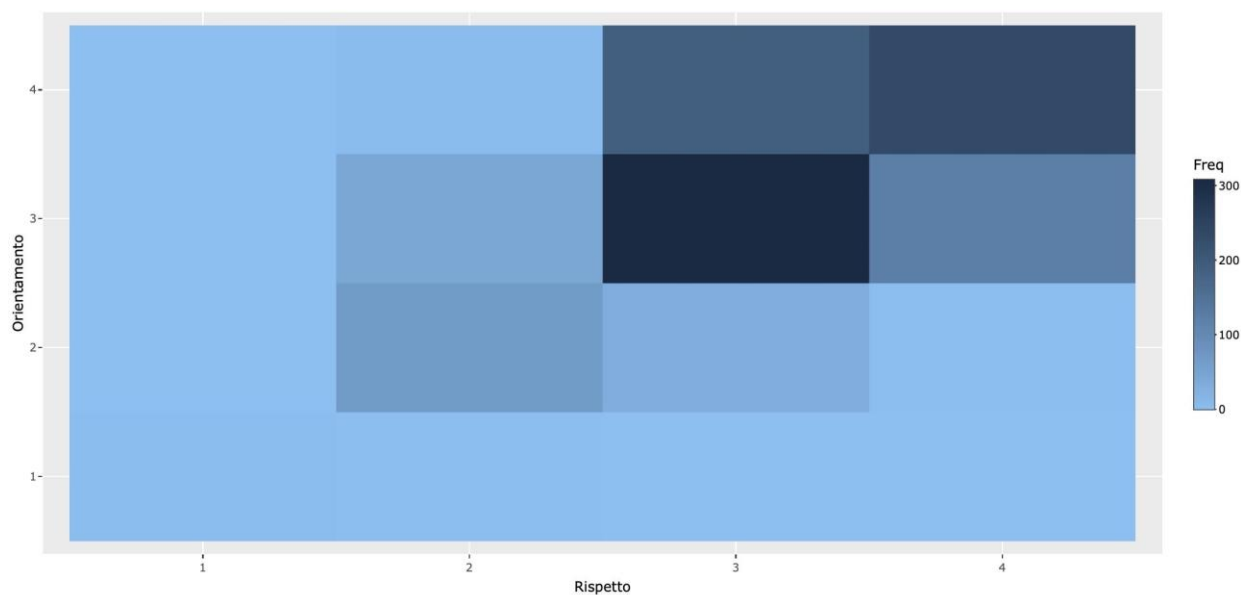
Fonte: elaborazione su dati ISPRA

L’ipotesi è che, nonostante la percezione del conflitto dimostrata dalla differenza dei pesi dati ai due item, entri in gioco un bias chiamato “effetto alone”: nel giudicare un individuo, veniamo influenzati da una

sua caratteristica positiva/negativa e tendiamo di riflesso a ritenere che egli abbia altre caratteristiche positive/negative che non sono in realtà collegate alla precedente (Nisbett & Wilson, 1977). Esistono diversi studi riguardanti l'effetto alone nella letteratura sulla formulazione di giudizi nella valutazione delle performance (ad esempio, è provato che un impiegato riceve un punteggio alto in ogni aspetto considerato, solo perché particolarmente bravo nello svolgere un singolo compito) (Belle et al., 2017). È anche ipotizzabile la possibilità di un errore di distribuzione (problema che si riscontra per tutti gli item e per tutte le valutazioni raccolte), problema noto in letteratura come “leniency error”, per il quale esiste una tendenza a dare valutazioni omogenee spostate verso l'alto (Kane et al., 1995; Ng et al., 2011). Tuttavia, la formulazione di giudizi omogenei su item formulati esplicitamente in contrapposizione tra loro è un fenomeno non affrontato in letteratura (figura 7).

Figura 7 - Tabelle di correlazione tra le valutazioni di Teamworking/Collaborazione intra ed extra funzionale e Rispetto delle procedure/Orientamento al risultato (2020)





Fonte: elaborazioni ISPRA

2.4 Ipotesi sul comportamento dei valutatori

Tenuto conto di quanto detto fin qui, vengono di seguito formulate ipotesi di interpretazione dei dati raccolti con l'utilizzo della scheda qui sopra analizzata, considerando i bias che sono stati evidenziati e che possono entrare in azione nell'affrontare ciascun item della scheda.

Verranno qui considerati sia i valutatori gerarchici che i valutatori in posizione di parità che valutano l'operato dei propri colleghi. Da notare che è possibile immaginare una distinzione tra le distorsioni cognitive del valutatore gerarchico e quelle del valutatore pari e dove sarà possibile, nelle ipotesi a seguire, tale distinzione sarà evidenziata.

Come dicevamo occorre, inoltre, distinguere tra i bias relativi agli item e quelli che derivano dalla particolare definizione che dell'item è stata data da chi ha redatto la scheda. La differenza tra i due tipi di bias è di fondamentale importanza al fine di immaginare possibili forme di intervento: nel caso dei bias derivanti dalla formulazione delle definizioni può infatti essere sufficiente un mutamento di definizione mentre nel caso dei bias relativi la scelta degli item, occorre una strategia di *debiasing* più sofisticata incidente sul processo di valutazione.

Item 1 - Il *teamworking* è definito come la "Capacità di collaborare con i membri del gruppo offrendo il proprio contributo e facendo frutto dell'apporto degli altri per il raggiungimento di uno scopo comune". Analizzando la formulazione delle possibili risposte, in relazione a questa definizione, dal punto di vista comportamentale possiamo ipotizzare che la scelta delle risposte si leghi alla formulazione proposta. Nel caso dell'item in questione, si può anche ipotizzare una diversa concezione di collaborazione all'interno del team nel caso di valutazione gerarchica e di valutazione tra pari (inoltre, si può ipotizzare che, quando queste concezioni divergono la combinazione tra valutazione gerarchica e paritaria può contribuire al *debiasing*). Conseguentemente, le nostre ipotesi sono le seguenti:

H1 - in caso di giudizio positivo, il valutatore propenderà verso la risposta 3, in quanto formulata esattamente come la domanda. Ciò a scapito della risposta 4, che differisce dal frame indotto dalla domanda introducendo termini e concetti non contenuti in essa (ovvero la validità del soggetto, la difficoltà del contesto e la possibile presenza di conflitti inter-relazionali).

H2 - se la concezione di collaborazione nel team del dirigente è diretta a integrare sistemi di competenze e non motivazionali mentre quella dei pari è concentrata sui profili motivazionali più che su quelli di competenza.

Item 2 – La definizione fornita per la qualità tecnica, cioè per la competenza dei soggetti nel loro specifico lavoro, è, a nostro avviso, priva di aspetti che potrebbero indurre bias. Anche i gradi di giudizio sembrano formulati in modo coerente con la definizione e senza particolari problematiche di sovrapposizione semantica. Di conseguenza, per questo item non possiamo ipotizzare bias indotti dalla formulazione. Analogamente, non si può a priori ipotizzare l'esistenza di bias generati da un incerto (e non trasparente) allineamento tra i pesi che a questo item danno i valutatori gerarchici e i valutatori tra pari.

Item 3 – Per quanto riguarda l'orientamento al risultato, è introdotta una scala eterogenea, che menziona in una (e una sola) delle definizioni un nuovo soggetto (l'utente) e un nuovo elemento (la soddisfazione dell'utente). In particolare, la scala da 1 a 4 dell'item nei punti 1, 2 e 3 fa unicamente riferimento al rispetto di regole e procedure. Inoltre, un dettaglio di rilievo relativo al presente quesito è l'introduzione delle attese dell'utente. Poiché risultato e utente possono essere legati positivamente, ma possono anche non esserlo (ad esempio un informatico, per ragioni di sicurezza, deve imporre policy restrittive sull'uso di internet, ciò che soddisfa la compliance normativa, ma non l'utente interno). In altri termini, non è chiaro in che modo sono legati questi concetti (orientamento al risultato, e aspettative dell'utente). È importante sottolineare che, nel caso di questo item, i bias dipendono non solo dalla particolare definizione che all'item è stata data, ma anche dall'item stesso e dall'insieme di valori e aspetti motivazionali che connota. In questo caso, la differenza nei tempi e nelle modalità di osservazione del comportamento, e cioè la differenza tra l'osservazione continua (che può essere effettuata dai pari) e quella discontinua (del dirigente) può essere molto rilevante nel processo di valutazione, poiché può comportare sia l'attivazione di bias diversi sia la necessità di strategie diverse per mitigare bias simili.

H3 – Nell'item “orientamento al risultato”, l'assenza del riferimento a regole e procedure nella risposta 4, coincidente con l'assenza di questi concetti nella definizione può indurre un giudizio positivo, facendo sì che l'opzione 4 venga scelta in modo sproporzionato (attribute framing)

H4 - il valutatore farà riferimento ad uno stereotipo o ad un'assunzione percepita, finendo per cedere all'attrattiva dell'item contenente il concetto di “aspettativa dell'utente”, (interpretando erroneamente il quadro generale in base ad una conclusione fatta a partire da un caso particolare) (bias di rappresentatività)

H5 - il valutatore gerarchico, più distante dalla pratica quotidiana e, quindi, meno in grado di distinguere l'attinenza delle procedure svolte con quelle richieste, ma allo stesso tempo “garante” del rispetto di queste ultime, potrà tendere ad essere più o meno severo nel valutare il rispetto delle procedure, rispetto al valutatore nella valutazione tra pari (ci aspettiamo, quindi, una divergenza sistematica e significativa tra le due risposte) (bias di corrispondenza).

Item 4 - Il rispetto delle procedure viene definito come “Conoscenza delle procedure interne e capacità di esecuzione delle stesse in modo da non causare ritardi, interruzione o problemi.” Come nel caso della “qualità tecnica” tutte le risposte paiono formulate coerentemente con tale definizione, per cui, anche in questo caso, non sono ipotizzabili bias legati alla formulazione, bensì si possono immaginare bias relativi ai due diversi metodi di valutazione (gerarchica e tra pari). In particolare, ipotizziamo

H6 – si potrà osservare una differenza sostanziale tra i due metodi di valutazione: poiché il quesito riguarda pratiche che possono direttamente riguardare lo svolgimento quotidiano delle attività e poiché in questo i colleghi sono soggetti più direttamente interessati, la valutazione tra pari sarà molto più severa rispetto alla valutazione gerarchica e questa disparità sarà osservabile, ancora una volta, dalla divergenza sistematica delle risposte. (errore fondamentale di attribuzione: simile al bias di corrispondenza ma espresso nella sua accezione legata alla tendenza a considerare il comportamento degli altri, specchio della capacità personale considerata inferiore alla propria; si veda, per un approfondimento, Gawronski, 2004).

Item 5 - La collaborazione intra/extra funzionale è definita come la “Capacità di relazionarsi e costruire rapporti produttivi con i colleghi della propria struttura e delle altre unità operative e con interlocutori esterni all'ente, assicurando un clima di collaborazione”. Ci sembra che questa definizione presenti aree di sovrapposizione con la definizione di *teamworking*: il concetto di “collaborazione intra/extra funzionale” porta al suo interno il concetto di *teamworking*. Mentre nella prima l'obiettivo è assicurare un clima di collaborazione, nel *teamworking* l'obiettivo è il raggiungimento di uno scopo comune. Guardando alla scala delle risposte da 1 a 4, sia nel *teamworking* sia nella collaborazione

intra/extra funzionale le definizioni sottendono che più si è collaborativi, migliori sono le probabilità di raggiungere uno scopo comune.

H7 - il valutatore tenderà a collegare le risposte date ai due item; in altri termini, la valutazione della capacità di creare un clima collaborativo sarà collegata alla capacità di raggiungere un obiettivo comune, esprimendo valutazioni correlate non necessariamente guidate da un'intenzionalità (bias di ancoraggio).

2.5 Indagine empirica per validare l'analisi comportamentale

L'obiettivo della fase sperimentale della ricerca è stato quello di identificare le possibili distorsioni cognitive che possono caratterizzare il processo di valutazione. L'analisi dei dati forniti da ISPRA ha rivelato che alcune valutazioni dei dirigenti avrebbero potuto essere influenzate dal modo in cui le schede di valutazione erano formulate (framing). L'indagine sperimentale ha indagato la relazione tra le distorsioni cognitive di coloro che hanno predisposto le schede e quelle di chi ha proceduto alla valutazione. In questo secondo ambito, inoltre, sono stati indagate sperimentalmente le possibili differenze tra valutazione gerarchica e tra pari nonché le possibili influenze nella valutazione dei dirigenti per il fatto che i dirigenti stessi sono a loro volta soggetti a valutazione.

Il disegno sperimentale si è focalizzato su alcuni temi oggetto della valutazione. In particolare, sul rapporto tra rispetto delle procedure ed orientamento al risultato dell'azione amministrativa.

L'analisi dei primi dati raccolti permette di formulare l'ipotesi di un bias di ancoraggio nel processo decisionale del valutatore indotto dal framing con cui sono state redatte le schede di valutazione. In particolare, la definizione dell'item "orientamento al risultato" e le risposte possibili presentano un'ambiguità riguardo la relazione tra questo aspetto e il "rispetto delle procedure". Tale ambiguità fa sì che i due item possano essere percepiti come in conflitto tra loro.

L'analisi preliminare dei dati raccolti dal Sistema ISPRA ha mostrato l'esistenza di una tendenza a dare valutazioni omogenee, spostate verso l'alto (problema che si riscontra per tutti gli item e per tutte le valutazioni raccolte), rendendo ipotizzabile l'intervento di un errore di distribuzione noto in letteratura come "*leniency error*" (Kane et al., 1995; Yildiz & Baltaci, 2009; Ng et al, 2011). Obiettivo di questa parte della ricerca empirica è stato dunque indagare le caratteristiche di questo fenomeno nel contesto considerato, raccogliendo dati sulle valutazioni dei singoli soggetti. In particolare, si è voluto comprendere la differenza tra valutazioni le cui conseguenze sono la irrogazione di sanzioni e l'assegnazione di premi e valutazioni il cui obiettivo è solo quello di acquisire informazioni senza che ne derivino conseguenze dirette sul piano economico. L'obiettivo è comprendere se la correlazione con premi e sanzioni pregiudichi la obiettività della valutazione orientandola ed in particolare impedendo che vi sia un'adeguata differenziazione tra i valutati. Più specificamente si è voluto testare l'ipotesi che una valutazione priva di conseguenze economiche produca una maggiore differenziazione e dunque esalti il merito rispetto ad una valutazione cui siano associate conseguenze immediate ed in particolare l'assegnazione di premi e l'irrogazione di sanzioni.

Le ipotesi alla base di questo studio sono le seguenti:

H1 – in un ambiente in cui le valutazioni servono esplicitamente a creare premi/punizioni si dovrebbe osservare un più forte leniency error rispetto ad un ambiente in cui la valutazione opera solo come strumento per il miglioramento delle performance.

H2 – in un ambiente in cui la valutazione gerarchica è esplicitamente utilizzata per distribuire premi/punizioni e quella tra pari solo come strumento di calibrazione, si dovrebbe osservare un leniency error maggiore nella valutazione gerarchica rispetto alla valutazione tra pari.

H3 – (social proof) quando il network di riferimento tende a dare dei giudizi uniformi il valutatore (sia pari che gerarchico) tenderà ad uniformare le proprie valutazioni.

Nel questionario proposto, sviluppato per comprendere la differenza tra valutazioni le cui conseguenze sono la irrogazione di sanzioni e l'assegnazione di premi e valutazioni il cui obiettivo è solo quello di acquisire informazioni senza che ne derivino conseguenze dirette, tale differenza viene colta immaginando due scenari:

- al rispondente viene chiesto di immedesimarsi nel ruolo di dirigente;
- al rispondente viene chiesto di immedesimarsi nel ruolo di collega

Un secondo e correlato profilo è quello relativo alle conseguenze della valutazione.

L'obiettivo è comprendere se la correlazione con premi e sanzioni pregiudichi la obiettività della valutazione orientandola ed in particolare impedendo che vi sia un'adeguata differenziazione tra i valutati. Più specificamente si intende testare l'ipotesi che una valutazione priva di conseguenze possa produrre una maggiore differenziazione e dunque esalti il merito rispetto ad una valutazione cui siano associate conseguenze immediate ed in particolare l'assegnazione di premi e l'irrogazione di sanzioni.

Integrando queste osservazioni nei due scenari sopra riportati, si possono ottenere le seguenti possibilità:

- che le valutazioni dei dirigenti abbiano conseguenze economiche dirette mentre quelle dei pari no;
- che sia le valutazioni dei dirigenti sia quelle dei pari incidano sull'assegnazione di premi e l'irrogazione di sanzioni;
- né le valutazioni dei dirigenti né quelle dei pari abbiano conseguenze perché il valore della valutazione è legato ai sistemi motivazionali e non agli incentivi economici.
- Un terzo profilo d'indagine concerne il valore delle norme sociali (Bicchieri, 2017). Si assume che vi sia una correlazione tra le modalità di valutazione e le norme sociali presenti nella organizzazione. Si vuole, quindi, indagare:
 - se le norme sociali influenzano la valutazione
 - come le norme sociali influenzano la valutazione
- In questo quadro si indagano due norme sociali:
 - quella concernente il ruolo della gerarchia. In che modo viene percepita la differenza tra superiore gerarchico e collega nel processo di valutazione;
 - quella concernente il conformismo ai valori dell'organizzazione e dunque al rapporto tra merito che dovrebbe risultare nella differenziazione dei giudizi e egualitarismo che esita nell' assenza di difformità.

Capitolo 3

PROPOSTA DI INTERVENTI COMPORTAMENTALI PER LA FORMULAZIONE DELLE SCHEDE DI VALUTAZIONE

Qui di seguito vengono presentati alcuni possibili interventi che, a vari livelli, possono essere implementati al fine di migliorare l'efficacia del nuovo sistema di valutazione, ideati in base a quanto discusso fin qui e sulla base delle rilevazioni empiriche fin qui effettuate (rimandiamo a tal proposito anche al documento specifico di analisi e interpretazione dei dati raccolti con il questionario pilota). Chiaramente tali proposte si distinguono in base a fattibilità e impatto e la decisione riguardo la loro effettiva implementazione dipenderà da queste due caratteristiche, da valutare accuratamente riflettendo sugli aspetti caratterizzanti l'organizzazione e il contesto esterno. Inoltre, si sottolinea che tali proposte potranno essere ulteriormente integrate e delineate (ed eventualmente rettificare, dove necessario), una volta elaborati e analizzati i dati del secondo questionario e dell'esperimento di laboratorio in via di somministrazione.

Per aumentare la consapevolezza e rendere la valutazione più forte di significato, si ritiene utile chiedere ai valutatori di assegnare un peso all'importanza di ciascun item da valutare, in relazione al tipo di lavoro da svolgere e, in modo astratto, rispetto al valutato. Inoltre, l'assegnazione tra tali pesi renderà la valutazione più aderente alle diverse componenti dell'organizzazione. Ci si aspetta che, ad esempio, alcuni item abbiano un peso maggiore se afferiscono al personale che svolge lavori più trasversali e che richiedono una maggiore interazione tra colleghi, mentre altri abbiano un peso maggiore in lavori più settoriali.

Nello specifico, si è deciso di provare a ridurre la distorsione indotta da un'assegnazione dei pesi preconstituita inserendo un sistema "*comply or explain*" dove si consente al singolo valutatore di indicare una diversità dei pesi attribuiti ai singoli item. Questo consente una maggiore autonomia dei singoli valutatori a discapito della uniformità del cui valore ovviamente tutti sono consapevoli. L'idea è, in altri termini, di rendere endogeni alcuni elementi cruciali, finora considerati in modo esogeno, eliminando anche il rischio della sostituibilità degli item di valutazione e inducendo nel valutatore un senso di responsabilità, contribuendo egli stesso alla definizione dello strumento di rilevazione delle valutazioni, tenendo in considerazione quegli elementi che il valutatore stesso considera importanti sia per quanto riguarda il valutato sia per la struttura organizzativa in cui agisce.

Per implementare questo intervento si è immaginata la possibilità di espressione dei propri pesi da parte del valutatore, in un momento precedente all'attività di valutazione stessa. Inoltre, sono previste due possibili modalità di espressione: una "privata", in cui il valutatore è chiamato ad esprimere il proprio parere riguardo ai pesi, semplicemente con la finalità di attivare in lui un più elevato livello di consapevolezza (per contrastare specificamente il bias di *mindlessness*); una "pubblica", in cui il valutatore è chiamato ad esprimere il peso attribuito agli item, sapendo che tale espressione di preferenza verrà resa pubblica (prevedendo, così, l'intervento dell'effetto della pubblicazione come una nuova variabile indipendente che va ad incidere sulla scelta del valutatore).

3.1 Alcune valutazioni preliminari sulle modifiche della scheda per il 2021

La scheda di valutazione ISPRA si fonda sul sistema di aspettative con un *range* da 1 a 4 che nella nuova scheda è stato portato ad un *range* da 1 a 7 (nell'implementazione effettiva per l'anno 2021 è stata introdotta considerando mezzi punti compresi tra 1 e 4, portando ad una scala del tipo 1, 1.5, 2, 2.5, 3, 3.5, 4. Ciò che viene chiesto al valutatore è di correlare la condotta del valutato alle aspettative

dell'organizzazione che implicitamente si assume siano anche quelle del valutatore. Dunque, non una valutazione 'obiettiva' scevra da riferimenti ad aspettative ma una valutazione basata sulla capacità del valutato di rispondere ad aspettative ex ante note al medesimo. Il valutato conosce la scheda e conosce il valore dell'aspettativa correlata all'item di valutazione.

Il processo di modifica della scheda di valutazione da parte di ISPRA è stato condiviso con il gruppo di ricerca. Alcune modifiche sono state apportate altre considerate in relazione alla valutazione relativa al 2021.

In questo paragrafo vengono indicati i suggerimenti proposti al fine di consentire un dialogo tra il gruppo di ricerca SNA e ISPRA.

Il sistema di valutazione in ISPRA è fondato sulle aspettative dell'organizzazione rispetto alla condotta dei valutati. Incorporare le aspettative consente di ridurre la *fictio* delle valutazioni obiettive che prescindono dai sistemi di credenze organizzative ed individuali.

L'esistenza di aspettative deve però essere obbiettivizzata in modo che siano conosciute e che siano chiare le relazioni tra aspettative dell'organizzazione, del valutatore ed aspettative del valutato.

La valutazione del direttore generale si basa sulle aspettative dello stesso rispetto ai dirigenti valutati. La valutazione dei dirigenti si basa sulle aspettative dell'organizzazione e su quelle proprie rispetto ai valutati. Le aspettative dei valutati riguardano sia l'organizzazione che i dirigenti valutatori.

In assenza di una valutazione bottom up il sistema di aspettative dei valutati non diviene mai oggetto di un processo di valutazione ma andrebbe incorporato nel sistema di valutazione da parte dei dirigenti. È evidente che quando il dirigente valuta si confronta con un sistema di aspettative dei valutati rispetto al quale può divergere o convergere. Tale sistema di aspettative ha una componente collettiva, il sistema di credenze della comunità dei valutati, ed una componente individuale, il sistema di credenze individuali di ciascun valutato. Il dirigente, in sede di valutazione, tiene implicitamente conto dell'aspettativa dei valutati riguardante sia le priorità attribuite ai singoli item da parte del dirigente sia il sistema complessivo di credenze (Cafaggi 2021).

In via di premessa occorre considerare che il sistema di valutazione è un sistema a cascata in cui le aspettative di valutatori e valutati sono correlate. Nell'indagare le possibili distorsioni cognitive dei dirigenti valutatori, è, dunque, necessario tenere conto del sistema di aspettative che i superiori del dirigente valutatore hanno e quelle dei dipendenti del dirigente valutatore.

Nel processo di valutazione il dirigente incorporerà l'aspettativa del proprio superiore sui profili organizzativi del gruppo soggetto a valutazione e quelle dei propri dipendenti sullo stesso tema. Scindere il processo valutativo dalle aspettative del dirigente verso l'alto e verso il basso della scala gerarchica oscurerebbe alcune delle fonti di processi cognitivi posti in essere nella valutazione da parte del dirigente. È solo facendo emergere il sistema di aspettative che è possibile interpretare correttamente i risultati del processo valutativo e dare spiegazione ad alcuni fenomeni quali quella della eccessiva uniformazione verso l'alto delle valutazioni così come emerge dai risultati del rapporto sulla valutazione relativo al 2019. (Cafaggi 2021).

Poste queste premesse di contesto veniamo ora all'esame della scheda di valutazione proposta ed alle possibili modifiche suggerite dall'impiego delle scienze comportamentali.

Si ritiene che le definizioni relative agli item andrebbero modificate in modo da rendere più chiaro e meno ambiguo il sistema di credenze di chi predispone le schede che ragionevolmente riflette la prospettiva di chi dirige l'organizzazione. Per questa ragione proponiamo integrazioni alle definizioni contenute precisando che la funzione della definizione è quella di ridurre l'occorrenza di bias derivanti dalla distonia tra il sistema di credenze organizzative circa team working, problem solving, orientamento al risultato etc. e quello dei singoli dirigenti valutatori.

Non è invece condivisibile la eliminazione della definizione del sistema di aspettative nella nuova scheda.

Le aspettative nella scheda attuale non vengono esplicitamente definite ma sono desumibili dalle definizioni degli item

La mancata definizione presterebbe il fianco alla critica di eccesso di soggettivismo ove, come nel caso della scheda modificata, le aspettative non vengano definite. Omettendo una definizione delle aspettative

del dirigente si assume che il dirigente valutatore abbia delle aspettative e che il valutato si comporti al di sotto, al livello, al di sopra di aspettative del dirigente senza che queste siano conosciute.

Ove si ritenga opportuno che il sistema di valutazione operi rispetto alle aspettative è fondamentale oggettivizzare le stesse evitando rischi di arbitrarietà definendo le aspettative e rendendole conoscibili ai valutati.

La definizione delle aspettative e la sua configurazione nei quattro livelli al di sotto, al livello, sopra, molto sopra concerne le aspettative dell'organizzazione. È dunque necessaria la definizione delle aspettative da parte dell'organizzazione. Tale definizione costringe il dirigente valutatore a confrontarsi con il sistema di aspettative dell'organizzazione quando venga posto in essere il processo valutativo. Oltre a raccomandare il reinserimento della definizione delle aspettative al fine di ridurre l'arbitrarietà ed i bias cognitivi ad essa associati solleviamo un'ulteriore questione.

Il problema riguarda la possibilità che le aspettative dell'organizzazione come cristallizzate nelle definizioni contenute nella scheda non coincidano con quelle del valutatore.

Cosa accade se la definizione di lavoro di squadra o di orientamento al risultato del valutatore diverge da quella definita nella scheda? Il valutatore si conformerà a quella data dall'organizzazione o procederà con la sua propria?

Ponendoci successivamente il problema della differenziazione dei valutati e del contrasto al *leniency error*, proponiamo di dare delle definizioni delle aspettative ma di modificarne il contenuto rispetto a quelle contenute nella precedente scheda di valutazione. In particolare, ciò che appare distorsivo è inserire nella definizione delle aspettative derivanti dalle correlazioni tra item. Dunque, si suggerisce di ripristinare la definizione delle aspettative eliminando le correlazioni tra item all'interno delle definizioni. Nella definizione dell'orientamento al risultato, ad esempio, la descrizione delle aspettative faceva riferimento esplicitamente al rispetto delle procedure identificando una correlazione tale per cui l'aspettativa di valore più elevato era quella del raggiungimento del risultato nel rispetto delle procedure. In verità mentre nei numeri 1,2,3 si fa riferimento al rispetto delle procedure nel numero 4 si omette.

Sotto il profilo comportamentale la possibilità di indurre bias nei valutatori cresce se le definizioni delle aspettative vengono presentate come correlazioni. Diversamente accadrebbe se le definizioni di ciascun item fossero costruite indipendentemente dagli altri item ed al valutatore fosse chiesto successivamente di valutare la correlazione.

Nel caso dell'orientamento al risultato questo andrebbe prima definito indipendentemente dal rispetto delle procedure e poi correlato a queste chiedendo di valutarne la capacità di conseguimento degli obiettivi con o senza il rispetto delle procedure.

Un secondo connesso problema riguarda la possibilità che le aspettative dei diversi dirigenti, che usano tutti la medesima scheda, divergano. Questo è assolutamente fisiologico potendo ciascun dirigente avere la propria idea sul modello organizzativo della propria unità e sui rapporti tra i diversi item valutativi. Allo stesso tempo è legittimo il tentativo dell'amministrazione di avere un processo di valutazione uniforme ed informato ad un sistema di credenze omogenee. Si tratta dunque, in sede di valutazione, di trovare un giusto compromesso tra esigenze di uniformità e sistemi di credenze individuali dei valutatori.

Cosa accade se la definizione delle aspettative differisce tra i dirigenti valutatori? L'eventuale difformità di aspettative ove prevalessero sulla definizione oggettiva potrebbe condurre ad una differenziazione tra valutazioni.

Questo fenomeno, che come sappiamo, in realtà, non ricorre, sarebbe in questo caso dipendente dalla differenza non esplicitata di credenze relative alle aspettative da parte dei singoli valutatori che, invece di attenersi al sistema di aspettative contenuto nella scheda, seguono il proprio.

Al fine di ridurre questo eventuale rischio, proponiamo che ISPRA chieda ai valutatori di indicare eventuali divergenze rispetto alla definizione delle aspettative contenute nella scheda. In tal modo emergerebbe con chiarezza il sistema di aspettative che è alla base della valutazione compiuta dai dirigenti.

Infine, il tema della correlazione tra item oggetto di valutazione

Le correlazioni tra item sono di fondamentale importanza per la individuazione di eventuali distorsioni. Nell'esempio dell'orientamento al risultato comprendere se dal punto di vista del valutatore sia rilevante che il conseguimento dell'obiettivo avvenga nel pieno rispetto delle procedure amministrative

interne o anche aggirando alcune di queste (ovviamente sempre nel rispetto di quanto consentito dalle norme) è molto significativo.

Tale indagine può svolgersi in due modi:

- a) correlando ex post le valutazioni dei due item tra i quali si vuole instaurare un rapporto oppure
- b) chiedendo esplicitamente di valutare un item in relazione ad un altro.

In conclusione, si propone di:

- Ripristinare la definizione delle aspettative, modificandola rispetto a quella contenuta nella precedente scheda (2020)
- Dare al valutatore l'opportunità di indicare se adotterebbe una definizione ed un peso diverso da quello proposto nella scheda al fine di far emergere possibili visioni differenti tra organizzazione e dirigenti valutatori
- Inserire la valutazione esplicita di alcune correlazioni come quella tra orientamento al risultato e rispetto delle procedure o tra sistema di relazioni e lavoro di squadra.

Tali modifiche consentirebbero di fare emergere visioni divergenti dei e tra i valutatori sulle caratteristiche del modello organizzativo in modo da tenerne conto nella interpretazione del processo valutativo

L'obiettivo di tali modifiche non è quello di eliminare la diversità di prospettive, fisiologica all'interno di un'organizzazione, ma di farla emergere in modo trasparente.

3.2 Contenere i bias creati dalla formulazione degli item

Con l'obiettivo di ridurre i bias del valutatore, proponiamo alcuni interventi sullo strumento di rilevazione delle valutazioni, a partire dalla definizione fornita per ciascun item nella scheda (Figura 1), che formuliamo con l'intento di ridurre l'ambiguità veicolata.

Prima di tutto, dunque, la nostra proposta è di riformulare le definizioni come segue:

1. *teamworking*: abilità/capacità di far parte di una squadra coesa che persegue un obiettivo comune pur nel quadro di possibili diverse visioni (anche in modalità agile);
2. orientamento al risultato: abilità/capacità di perseguire, con costanza e determinazione, obiettivi di carattere professionale o personale coerenti con il piano di performance organizzativo (anche in modalità agile);
3. qualità tecnica: insieme di competenza e esperienza acquisita in un determinato settore o materia;
4. rispetto delle procedure: abilità/capacità di condurre un lavoro nel rispetto dei vincoli normativi esistenti (per vincolo normativo si intende sia la legislazione nazionale o internazionale, sia le procedure amministrative interne) (anche in modalità agile);
5. collaborazione intra/extra funzionale: abilità/capacità di relazionarsi e costruire rapporti produttivi con i colleghi della propria struttura e delle altre unità operative e con interlocutori esterni all'Ente (anche in modalità agile).

In secondo luogo, rispetto alle correlazioni, come anticipato, si suggerisce di inserire alcuni item che indaghino la relazione tra item di valutazione. È possibile anche tener conto della correlazione tra i diversi item, considerando che (a) l'item 5 è correlato all'item 1, poiché l'attitudine alla collaborazione facilita la capacità di *teamworking* e (b) l'item 2 è correlato all'item 4: siccome è necessario rispettare i vincoli normativi esistenti, l'aspettativa dell'orientamento al risultato deve tener conto di tale "limitazione" (in altri termini, il dirigente deve assegnare obiettivi raggiungibili che presuppongono il rispetto dei vincoli normativi). Nel caso venga accolta la proposta di inserire le correlazioni come item di valutazioni è necessario indicare anche le aspettative con la definizione delle stesse per le ragioni indicate in precedenza.

Ove non si ritenga opportuno inserire quali autonomi item di valutazione le correlazioni, queste possono rappresentare elementi di controllo nella coerenza della valutazione. Ad esempio, ci si aspetta che chi ha scarsa capacità di collaborare abbia anche un basso punteggio sul team working, così come il raggiungimento dei risultati non dovrebbe essere posto come elemento antitetico al rispetto delle procedure.

Figura 1 – Scheda di valutazione per il personale I-VIII (rev. 2021)

SCHEDA DI VALUTAZIONE PER IL PERSONALE I-VIII

Unità Operativa	Valutato	Valutatore					
Criteria di valutazione			Peso	Valutazione		Punteggio	
Performance organizzativa della struttura organizzativa di afferenza ex Piano della <i>performance</i> (criterio a.)			85%				
Comportamenti agiti e competenze di ruolo (criterio c.)			15%				
Comportamenti agiti e competenze di ruolo (criterio c.)							
Descrizione competenze e scala di valutazione							
			<i>al di sotto delle aspettative</i> 1.0 – 1.5	<i>in linea con le aspettative</i> 2.0 – 2.5	<i>sopra le aspettative</i> 3.0 – 3.5	<i>Molto oltre le aspettative</i> 4.0	Peso
Teamworking: Abilità/Capacità di perseguire il raggiungimento degli obiettivi in maniera coordinata ed in stretta collaborazione con i colleghi della struttura organizzativa di appartenenza, pur nel quadro di possibili diverse visioni nel (anche in modalità agile)							20%
Qualità tecnica: Capacità di contribuire, con competenza tecnica e professionalità, al raggiungimento degli obiettivi della struttura organizzativa di appartenenza e capacità di problem solving (anche in modalità agile)							20%
Orientamento al risultato: Capacità di orientare la propria attività alla prioritaria realizzazione degli obiettivi della struttura organizzativa di appartenenza							20%
Rispetto delle procedure: Abilità/capacità di condurre e realizzare la propria attività nel rispetto delle indicazioni operative definite (circolari, linee guida, procedure amministrative) (anche in modalità agile)							20%
Collaborazione intra ed extra funzionale: Abilità/Capacità, nell'esercizio dei compiti assegnati, di relazionarsi e costruire rapporti produttivi non solo con i colleghi della propria struttura ma anche con le altre unità operative e con gli interlocutori esterni all'Ente (anche in modalità agile)							20%
<i>Gli esiti della valutazione sopra riepilogati sono stati comunicati in un "colloquio di feedback" contestualmente alla consegna della presente scheda al valutato.</i>							
<i>Eventuali osservazioni da parte del valutato</i>							
Luogo e data	Firma Valutatore			Firma Valutato			

Fonte: ISPRA

3.3 Indurre la differenziazione modificando la scala

Uno degli obiettivi perseguiti da ISPRA è quello di indurre i dirigenti a adottare valutazioni differenziate riducendo il fenomeno dell'uniformità verso l'alto. Le valutazioni effettuate nel 2020 presentano un elevato grado di uniformità verso l'alto (come mostrato dalla Figura 1).

Come detto, il presupposto teorico del nostro approccio è l'idea di Kahneman (2002) che esistano due "sistemi" all'interno della mente che, una volta attivati, prendono il "controllo" delle nostre decisioni: il Sistema 1, automatico, impulsivo e non soggetto ad elaborazione di alto livello, e il Sistema 2, relativo al ragionamento accurato e ad un'elaborazione che consuma tempo ed energia. Il primo permette di decidere velocemente ma può indurci in errore (legato al ragionamento euristico), il secondo è accurato, ma molto lento. Tenendo presenti queste considerazioni e riportandole all'analisi di ciò che può accadere in fase di valutazione, possiamo dire che l'adozione di una scala da 1 a 10, che sostituisca la scala da 1 a 4, può servire a spingere verso una desiderata differenziazione dei giudizi. Tuttavia, poiché questo "sgranamento" potrebbe indurre sia il valutatore che il valutato a percepire la valutazione non come la misura soggettiva delle attese comportamentali ma un giudizio complessivo sull'individuo, per recuperare parzialmente l'intenzione di attivare i meccanismi dell'attenzione, una possibilità è quella di spostare la linea della "sufficienza" ad un valore inferiore al 6, ottenendo il duplice effetto di ottenere un numero più elevato di giudizi positivi e ammortizzare l'"effetto pagella", costringendo il valutatore ad uscire dall'automatismo che la scala da 1 a 10 potrebbe indurre ("sistema 1"), attivando il "sistema 2" di ragionamento.

Il problema di una modifica di questo tipo, tuttavia, è che difficilmente potrà effettivamente ridurre la propensione ad attivare automatismi appresi in un sistema di valutazione così ampiamente utilizzato come la valutazione da 1 a 10. La proposta che si potrebbe prendere in considerazione in questo caso è, da una parte, di introdurre una scala del tutto inusuale, che permetta (a) al sistema 2 di attivarsi (non potendo la mente del valutatore richiamare facilmente l'automatismo "pagella"), (b) un valore medio su cui soffermarsi per scegliere se valutare più in alto o più in basso di esso (questo potrebbe essere importante per rendere più consapevole l'eventuale scelta di "essere clementi" o "essere severi"), Tabella 1.

Tabella 1 – Esempio di formulazione delle definizioni dell'item e dei gradi di giudizio e scala da 1 a 7.

Teamworking:	1	2	3	4	5	6	7
abilità/capacità di far parte di una squadra affiatata che persegue un obiettivo comune (anche in modalità agile)	Inefficacemente capace	Poco capace	Meno capace della norma	Normalmente capace	Più capace della norma	Molto capace	Eccezionalmente capace

3.4 Riduzione del bias of mindlessness con l'espressione di pesi

Un ultimo elemento contenuto nella scheda è il peso attribuito ai diversi item. Come è possibile notare, nella scheda utilizzata nella valutazione 2020 il peso dei cinque item nella valutazione complessiva era considerato esattamente identico per ciascun item (20%).

Nella nuova scheda, vengono proposti pesi diversi calcolati preventivamente in base all'indagine empirica condotta (tenendo conto dei pesi attribuiti dai dirigenti Ispra e dai pesi deducibili dai dati del questionario pilota). In particolare, se si considerano i pesi associati a ciascun item da parte dei dirigenti ISPRA, come si evince dalla Figura 1, si può osservare che la qualità tecnica ha un peso nettamente preponderante rispetto agli altri, seguito da *teamworking* e collaborazione intra/extra funzionale, percepite allo stesso modo e che presentano, al momento di questa analisi, anche aree di sovrapposizione nella definizione.

Il peso dell'orientamento al risultato segue molto da vicino i due precedenti. Da ultimo, nettamente distanziato, si trova il peso associato al rispetto delle procedure, percepito, anche nella definizione, come un ostacolo al raggiungimento dei risultati.

La riduzione della distorsione indotta da una scheda di valutazione precostituita si può realizzare inserendo un sistema *agree or explain* dove si consente al singolo valutatore di indicare una diversità dei pesi attribuiti ai singoli item. Questo consente una maggiore autonomia dei singoli valutatori a discapito della uniformità del cui valore ovviamente tutti sono consapevoli. L'idea è, in altri termini, di rendere endogeni alcuni elementi cruciali, finora esogenamente imposti, eliminando anche il rischio della sostituibilità degli item di valutazione e inducendo nel valutatore un senso di responsabilità, contribuendo egli stesso alla definizione dello strumento di rilevazione delle valutazioni, tenendo in considerazione quegli elementi che il valutatore stesso considera importanti sia per quanto riguarda il valutato che la struttura organizzativa in cui si agisce. Per implementare questo intervento si può immaginare la possibilità di espressione dei propri pesi da parte del valutatore, in un momento precedente all'attività di valutazione stessa. Questa espressione potrà essere effettuata tramite una scheda attraverso la quale indicare il peso da attribuire a ciascun item valutativo, analoga a quella sottoposta ai dirigenti Ispra ex-post nel 2020, oppure, in alternativa, analoga a quella somministrata nella domanda 1 del questionario sviluppato nel corso della ricerca. Inoltre, si possono prevedere due modalità di espressione: una "privata", in cui il valutatore è chiamato ad esprimere il proprio parere riguardo ai pesi, semplicemente con la finalità di attivare in lui un meccanismo di ragionamento (per contrastare il bias di mindlessness); una "pubblica", in cui il valutatore è chiamato ad esprimere il peso attribuito agli item, sapendo che tale espressione di preferenza verrà resa pubblica (prevedendo, così, l'intervento dell'effetto della pubblicazione come una nuova variabile indipendente che va ad incidere sulla scelta del valutatore).

L'idea è di approfondire la differenza sia nell'implementazione sia nelle conseguenze ipotizzabili e, cioè, nei comportamenti osservabili. Probabilmente un ruolo molto importante in questa dinamica sarà svolto dal "*social desirability bias*" (Krumpal, 2013): manipolando le variabili che possono indurre la sua attivazione possiamo ottenere comportamenti diversi (scegliendo di intervenire con strategie di sfruttamento del bias alternativamente a strategie di *debiasing*). L'efficacia dell'utilizzo di questo intervento comportamentale, per il momento solo parzialmente implementato, potrà essere misurata non appena i dati relativi la valutazione 2021 saranno resi disponibili: si potrà infatti procedere ad un confronto tra la valutazione effettuata nell'anno precedente in base ai pesi espressi solo ex-post e che abbiamo riportato nella sezione dedicata e la valutazione effettuata nel 2021 con l'espressione dei pesi in concomitanza con la fase di valutazione.

3.5 Primi risultati in base alla preliminare elaborazione dei dati raccolti nel 2021

Premettendo che l'analisi dei dati raccolti con la nuova scheda non è stata ancora completata, è possibile osservare alcuni primi risultati, grazie ad una preliminare elaborazione effettuata su questi dati al fine di poter avere una comparazione con i dati raccolti nel 2020. In particolare, si è proceduto all'osservazione della deviazione standard e della media dei giudizi, per poter comprendere se e in che modo ha inciso il percorso formativo sul problema della differenziazione delle valutazioni.

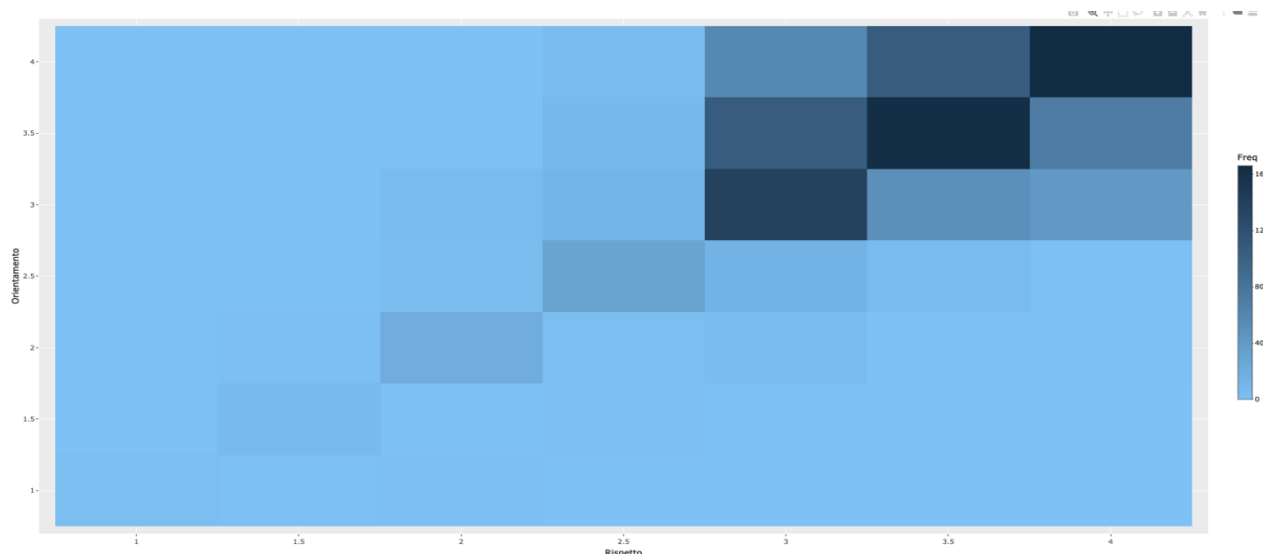
A tal proposito, si richiama la Tabella 2, delle valutazioni generali raccolte in aggregato, per tutti i soggetti valutati che quest'anno sono aumentati da 1016 a 1043 (data la scarsa significatività della differenza tra i due numeri, rimangono facilmente comparabili anche senza fare ricorso alle percentuali).

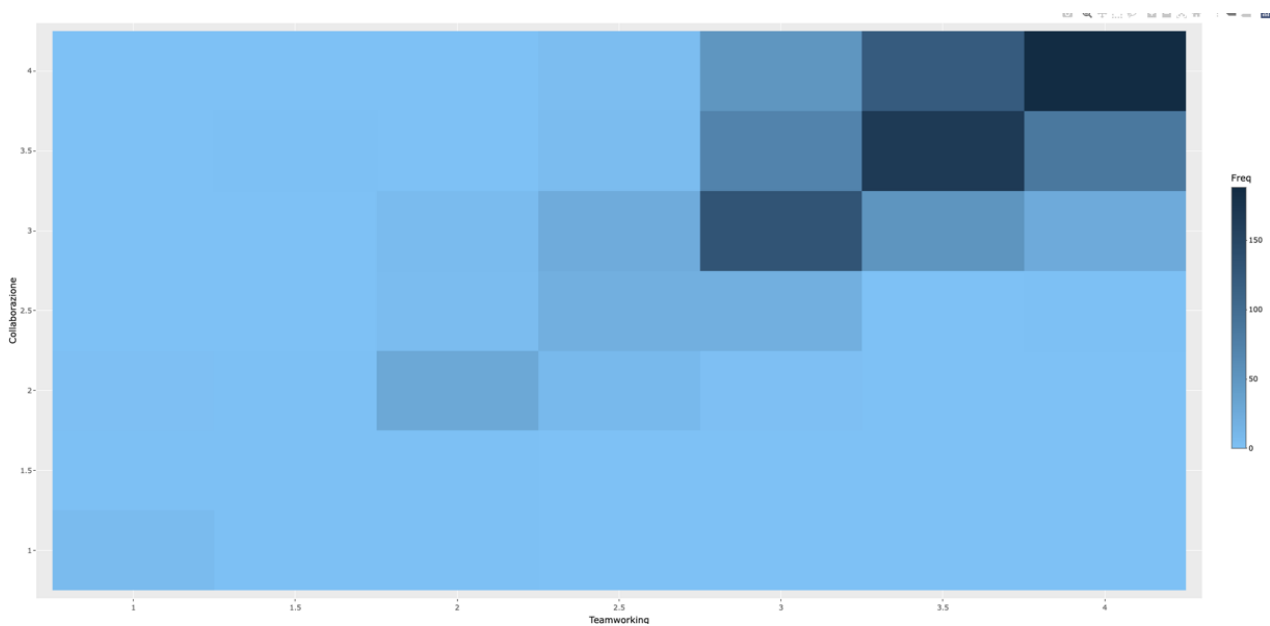
Tabella 2 - Dati aggregati sulle valutazioni generali (2021)

	X=0	0<x≤0.25	0.25<x≤0,5	x>0,5
Deviazione standard	166	186	585	106
Media dei giudizi	$x \leq 2$	$2 < x \leq 3$	$3 < x \leq 3.5$	$x > 3.5$
	32	157	371	560

Dai valori sopra riportati (Tabella 4), sebbene non si possa evincere un'effettiva risoluzione del problema, si può osservare un dato interessante, rispetto a quanto raccolto nella valutazione effettuata nell'anno precedente: quelli che danno tutti lo stesso valore sono diminuiti molto. È possibile osservare, dunque, che, sebbene la dispersione dei valori non sia cambiata significativamente, l'effetto alone sembra essere stato di fatto mitigato. Anche le valutazioni in cui la deviazione standard sia uguale a 0 appaiono diminuite e le deviazioni standard appaiono più contenute mediamente (deviazione standard media è 0,3 nei nuovi dati e 0,4 nei vecchi) e ci sono più deviazioni standard tra 0,25 e 0,5 e una quantità minore di deviazioni standard maggiori di 0,5. Da ciò è ipotizzabile una maggiore attenzione da parte dei valutatori agli item di valutazione e un affievolimento, quindi, dell'incidenza del mindlessness bias. Non è possibile osservare, invece, un effettivo miglioramento per quanto riguarda il contrasto al leniency error, poiché la media dei giudizi espressi appare essere sempre molto alta (sebbene vada segnalato il fatto che la metà circa è sotto 3,5). Infine, riportiamo di seguito (Figura 2) i grafici raffiguranti le correlazioni dei giudizi espressi negli item oggetto di attenzione per la nostra analisi comportamentale della valutazione effettuata nel 2020, relativamente ai dati raccolti nel 2021.

Figura 2 - Tabelle di correlazione tra le valutazioni di *Teamworking/Collaborazione intra ed extra funzionale* e *Rispetto delle procedure/Orientamento al risultato* (2021)





Fonte: elaborazioni ISPRA

Da quanto si può osservare, sembra che la correlazione tra “rispetto delle procedure” e “orientamento al risultato” si sia ulteriormente intensificata e, in modo analogo, sebbene meno marcato, si sia evoluta la correlazione tra *teamworking* e “collaborazione intra-extra funzionale”. Per dare conto di questi risultati, oltre alla possibilità di effettuare ulteriori approfondimenti, è probabilmente utile analizzare le riflessioni derivate dal confronto effettuato tramite il primo focus group formativo, che ha coinvolto buona parte dei valutatori, riportate nella sezione che segue.

Capitolo 4

ESITI DEL FOCUS GROUP FORMATIVO

L'impostazione data all'intero lavoro, a partire dalle considerazioni sugli effettivi obiettivi del Sistema di valutazione, ha caratterizzato anche la direzione del focus group: non ci si è concentrati sulla valutazione per l'assegnazione dei premi e incentivi ma come strumento di carattere organizzativo, per governare le possibili differenze legate alla diversità di visione sul modello organizzativo

Il punto cruciale è che i livelli gerarchici possono avere visioni e sistemi di aspettative diversi e che è importante fare emergere queste aspettative per governare e non occultare la diversità, al fine di attuare un processo di acquisizione di consapevolezza.

La scheda di valutazione predisposta da ISPRA individua nelle aspettative un elemento importante, per cui ragionare sul significato delle aspettative come parametro di misurazione è fondamentale per capire esattamente quali aspettative vengono prese in considerazione e come esse vengono incorporate negli indicatori. Quello che emerge dall'analisi è che convergono, nella valutazione fatta dai dirigenti rispetto ai collaboratori e ai dipendenti, almeno tre sistemi di aspettative che incidono sulla valutazione:

1. le aspettative di chi redige la scheda di valutazione (della dirigenza) rispetto a chi valuta;
2. le aspettative dei valutatori rispetto ai valutati, ovvero il rapporto tra i diversi dirigenti, cioè le aspettative che ci si aspetta gli altri dirigenti abbiano rispetto alla propria unità;
3. le aspettative dei valutati.

Ciononostante, esplicitamente evidenziato nella scheda di valutazione è solo l'aspettativa del valutatore rispetto al comportamento del valutato, mentre non emergono gli altri due sistemi. Il lavoro svolto nel progetto è diretto a discutere il rilievo dell'insieme dei sistemi di aspettative e nel caso di loro divergenza comprendere in che modo essa possa essere governata nel processo di valutazione.

4.1 Obiettivo del focus group: comprendere il ruolo delle aspettative

Il focus group, realizzato il 4 giugno 2021, è stato finalizzato proprio ad una discussione che iniziasse a portare alla luce il fatto che quando si procede alla valutazione, si prendono in considerazione, consapevolmente o meno, non solo le aspettative della direzione generale, ma anche quelle dei colleghi pari e dei valutati.

L'obiettivo finale è di capire cosa succede se divergono i sistemi valoriali tra diversi livelli gerarchici rispetto ciò che è importante valutare.

Sotto il profilo delle possibili divergenze ci sono almeno due questioni da affrontare:

1. La misura in cui sono condivise dai dirigenti le priorità attribuite ai criteri di valutazione stabiliti dalla direzione nella scheda di valutazione oppure si ritiene, per ragioni legate alla soggettiva visione dell'organizzazione o per ragioni legate alla specificità dell'area che si dirige, che altri criteri sarebbero più rilevanti. C'è in questo caso un trade-off tra la necessità di avere un sistema di valutazione uniforme e la necessità di avere delle aree più specializzate nelle quali si introducono elementi di valutazione specifici (es. aree prevalentemente amministrative rispetto a quelle più tecniche possono richiedere tipi di competenze e capacità e quindi richiedere nell'ambito di parametri di valutazione l'introduzione di profili differenziati). Il focus group ha evidenziato le opportunità derivanti dalla partecipazione dei valutatori al disegno della valutazione con suggerimenti e indicazioni rilevanti sia sul sistema di aspettative che su quello dei pesi attribuiti agli indicatori. Emerge, dunque, la possibilità di favorire la compartecipazione al disegno (che sia ex-ante o ex-post) dei dirigenti che andranno ad effettuare la valutazione. In particolare, tale

partecipazione dovrebbe consistere non solo in un impiego ragionato dei criteri esistenti, ma anche in una riflessione su ciò che manca nella scheda di valutazione e che andrebbe integrato (o in via generale o a livello delle singole specificità, evitando che questa differenziazione interferisca sulla necessaria uniformità ma, allo stesso tempo, si conformi al sistema di valori e visioni che ciascuno ha).

2. Legato a ciò che è stato già effettuato durante la prima fase della ricerca in corso, c'è poi il discorso dei pesi attribuiti agli item contenuti nella scheda di valutazione. Ciò che abbiamo rilevato grazie a quella attribuzione di pesi è un sistema di valori che indica che la qualità tecnica viene decisamente al primo posto, mentre il rispetto delle procedure va in coda. Un'altra cosa che risulta dai pesi raccolti è anche una differenziazione tra i pesi attribuiti; per osservarla abbiamo calcolato un indice di dispersione del dato, che mostra quanto siano diversi i pesi attribuiti a ciascun item all'interno del cluster utilizzato per la raccolta. Sebbene i dati raccolti siano molto pochi, si può osservare che la deviazione più grande si osserva nei pesi attribuiti al teamworking (è quasi ipotizzabile una polarizzazione: alcuni lo considerano un item molto pesante, altri molto basso), mentre la più piccola differenziazione è osservabile riguardo al peso della qualità tecnica (peso molto alto espresso da tutto il campione). La discussione nel focus group dovrebbe aiutare a consolidare ciò che si può dire su questi pesi, grazie alla partecipazione di un numero più ampio di soggetti interessati.

In relazione sia alla differenziazione dei sistemi di aspettative che a quella dei pesi, è importante sottolineare che un modello organizzativo che funziona non oscura le differenze, ma le fa emergere e le governa. Ciò che ci proponiamo non è portare le varie visioni a convergenza, ma di essere in grado, come sistema organizzativo, di fare emergere le diversità e gestirle. In questo quadro è importante tornare ad una proposta fatta lo scorso anno e temporaneamente accantonata: affinché questi sistemi di aspettative che guidano il processo di valutazione e le differenze di visioni possano emergere ed essere temperate e tradotte in utili indicazioni, è importante combinare la valutazione che i dirigenti fanno dei propri collaboratori e un "autovalutazione" che i collaboratori fanno della propria comunità. In altri termini, per fare emergere il sistema di aspettative dell'organizzazione nella sua totalità, è di fondamentale importanza affiancare alla valutazione che i dirigenti fanno dei propri collaboratori, anche la valutazione che i valutati fanno rispetto a sé stessi e rispetto i propri colleghi pari grado.

4.2 Realizzazione del focus group: aspettative rivelate e partecipazione alla progettazione

Durante il focus group si sono discusse le differenze introdotte nella scheda di valutazione attuale, relative alla definizione degli item di valutazione, elemento fondamentale proprio correlato alle aspettative, poiché tanto minore è la possibilità di avere interpretazioni ambigue degli item di valutazione, tanto maggiore è la consapevolezza del processo che si svolge (i.e. se ci sono delle valutazioni non necessariamente condivise, ma chiare, diventa conoscenza comune ciò che la direzione vuole che venga valutato). Inoltre, è stata data ai valutatori la possibilità di esplicitare la propria interpretazione, per far emergere elementi specifici da integrare rispetto a quelle definizioni o, anche, agli item stessi.

A conclusione dei lavori, oltre ad essere emersi i diversi sistemi di aspettative, si è potuto riprendere il tema dei pesi attribuiti dai singoli valutatori agli item di valutazione, permettendo considerazioni importanti sulla visione strategica dell'organizzazione e su alcune diverse visioni che dai dati precedentemente raccolti apparivano sono parzialmente. In particolare, si sono raccolti alcuni elementi fondamentali per capire come la differenza che risulta dai pesi raccolti è espressiva della differenza dei diversi modi di vedere i diversi item in base al tipo di attività che si svolge: fatti oggettivi che caratterizzano il modello di lavoro di ciascuno conducono a pesi diversi.

Grazie a questo esercizio si è iniziato a sviluppare un quadro chiaro delle differenze non solo tra la vostra visione intesa come visione complessiva della dirigenza Ispra rispetto alla visione della direzione

generale, ma anche delle divergenze all'interno del gruppo di dirigenti. L'obiettivo delle modifiche che saranno apportate al sistema di valutazione in generale (nelle prossime tornate) e alle schede in particolare (introducendo già da ora alcuni correttivi in base alle osservazioni emerse e che sottolineano alcune convergenze su cui è giusto focalizzarsi) è di dare traccia delle differenze distintive riscontrate e che si riscontreranno.

In seguito, dunque, tenendo presente il fatto che il sistema di valutazione è un sistema molto complesso che deve tener conto da una parte della differenziazione delle diversità esistenti dall'altra della necessità di essere un'organizzazione omogenea, per migliorare il percorso (ricordiamo che sono sistemi da adattare anno dopo anno), si valuterà la misura in cui le differenze evidenziate possono essere incorporate rivedendo la proporzione tra una parte uniforme e una parte destinata alla valutazione delle singole direzioni o dei singoli dipartimenti.

4.3 Risultati dei lavori di gruppo e proposte per il processo di valutazione 2021

Il lavoro condotto durante i focus group ha dato la possibilità di delineare alcune modifiche da apportare concretamente al sistema di misurazione e valutazione delle performance attualmente utilizzato in ISPRA.

In particolare, si è sottolineata, prima di tutto, l'opportunità di non aumentare per il momento il numero degli item in un'operazione di razionalizzazione, sebbene la risposta più immediata sarebbe quella di aumentarli per aumentare la differenziazione nelle valutazioni, al fine di non rischiare di diventare uno strumento solamente funzionale alla differenziazione (finendo per valutare, invece, profili analoghi).

Per quanto riguarda i pesi sugli attuali item, si è visto che laddove c'è un'attività più gestionale amministrativa rispetto delle procedure e orientamento al risultato assumono i primi posti, laddove invece c'è un'attività più tecnico scientifica assumono importanza la qualità tecnica e la capacità di collaborare.

Sono, inoltre, state evidenziate criticità legate alla percezione dei valutati. Ciò che è emerso è che la maggior parte delle obiezioni e delle critiche dei valutati è che manca un'esplicita definizione degli obiettivi personali dei valutatori. Chiaramente la risposta a queste critiche è che la valutazione avviene riguardo ciò che viene fatto dai singoli individui all'interno della struttura. Il suggerimento è di aggiungere negli item, quando si parla di obiettivi attribuiti, "alla struttura di appartenenza", per esplicitare questo aspetto.

Si è discusso di come, anche per una serie di confronti, sembra che gli item finiscano per essere un particolare di una persona in generale, non solo del lavoratore. Si tratta di item che possono essere ricondotti ad alcune macro-categorie di valutazione dell'individuo, dei profili più generali, che includono qualità tecniche (macro-categoria che viene, unico caso, rappresentata interamente da un singolo item) relazionali, etiche e personali (in quest'ultima macro-categoria rientrerebbe l'autonomia, ma anche altre, come l'ottimismo, ad esempio). Valutare le macro-categorie piuttosto che singole accezioni potrebbe essere più semplice per arrivare ad una valutazione consapevole.

Passando alla descrizione dei singoli item di valutazione si sono evidenziati diversi importanti aspetti:

- si è evidenziata la natura limitrofa e la parziale convergenza del concetto di *teammworking* e ciò che viene definito "collaborazione extra-funzionale": vanno entrambi ad indagare un aspetto comportamentale, cioè la capacità di lavorare in gruppo, che viene vista sotto due aspetti diversi: all'interno del team o con soggetti che non fanno parte del proprio team. Ciononostante, andando ad indicare comunque aspetti diversi ed essendo rilevante per un istituto che vuole enfatizzare le attività trasversali tra diversi centri, nell'ipotesi di riorganizzazione, sembra corretto mantenerle separate;
- è stato individuato un aspetto che meriterebbe un'attenzione particolare: l'autonomia (importantissimo per un dirigente da riscontrare nei propri collaboratori). L'autonomia può essere definita come la capacità propositiva diretta (cioè la capacità di proporre qualcosa senza stimoli endogeni) o indiretta (cioè la capacità di proporre soluzioni all'insorgere di un problema).

- si è sottolineata la correlazione tra *problem solving* ed innovazione stabile di medio lungo periodo nel modello organizzativo. Centralità del concetto di “*problem solving*” associato all’elemento di innovatività: non solo risolvere problemi, ma proporre soluzioni che possano essere applicate anche in altre situazioni;
- si è preso atto della difficoltà di modificare nella scheda di valutazione gli item attuali, attraverso un accordo tra valutatori, validato dalla direzione ISPRA, è possibile indicare in quale dei profili questi aspetti possano essere incorporati già nel processo di valutazione nell’anno 2021. Ad esempio, in “orientamento al risultato”, l’individuazione di modalità innovative di risolvere problemi nella vecchia scheda era considerata in un altro item di valutazione, la “qualità tecnica”. Quindi, è sicuramente presente e importante e si può provare a reintrodurre nella scheda semplificata, sia nella qualità tecnica che nell’orientamento al risultato, nella doppia accezione. Considerando che può trovare integrazione in una o nell’altro in situazioni e ambiti diversi: in ambito amministrativo può essere ricondotto all’orientamento al risultato, in ambito tecnico-scientifico può essere ricondotto alla qualità tecnica. Essendo due aspetti della stessa medaglia, crearne un item apposito potrebbe essere la soluzione migliore, ma sarà possibile solo in futuro. In questa tornata si può inserire, invece, come interpretazione;
- è stata sottolineata l’importanza sostanziale ma la non chiarezza dell’item “rispetto delle procedure”: così come formulata ora, la definizione non permette di disambiguare se si fa riferimento a procedure interne alla singola unità o a procedure strettamente vincolanti di tipo legislativo o, ancora, a prassi consolidate e non codificate. Andrebbe, quindi, probabilmente esplicitato, per dare peso alle prassi interne (all’interno della scheda viene riportato vincolo normativo per cui sembra molto stringente). L’interpretazione autentica è che si tratta di un elemento che va oltre la norma di legge e si intende come rispetto delle procedure interne così come interpretate dal singolo dirigente. Tuttavia, si tratta di procedure definite e non semplicemente agite. Sarà probabilmente necessario definire in modo più preciso a che tipo di procedure si fa riferimento: non le leggi, piuttosto, da una parte, le linee guida o indicazioni operative che vengono date da strutture sovraordinate deputate a fornire queste indicazioni di merito e, dall’altra, per quanto riguarda la prospettiva interna, procedure fissate per operare al meglio.

Capitolo 5

PROPOSTA DI INTERVENTI COMPORTAMENTALI SULLA STRUTTURAZIONE DEL PROCESSO DI VALUTAZIONE

Come si è detto¹, le proposte fin qui delineate sono da intendersi come modifiche puntuali al sistema di valutazione, tramite interventi comportamentali mirati all'eliminazione di possibili bias grazie alla riformulazione delle schede, da una parte, e all'introduzione di metodi esplicitazione delle aspettative, dall'altra. Si possono in qualche modo considerare proposte di implementazione di strategie comportamentali da modulare e adattare nel tempo, ma i cui effetti possono essere visibili in un periodo relativamente breve.

Tuttavia, è possibile immaginare interventi comportamentali più complessi e da applicare in un medio periodo per un cambiamento di alcuni passaggi dell'intero processo di valutazione. A seguire presentiamo due possibili percorsi in questa direzione

1.1 Riduzione del bias di mindlessness con espressione di preferenze su item incrociati

In seguito alla incongruenza riscontrata riguardo la correlazioni tra le risposte di item che avrebbero dovuto essere considerati in contrasto e che sono stati invece valutati congiuntamente (si veda paragrafo 3.2), potrebbero essere introdotte schede preliminari da sottoporre ai valutatori in cui essi siano chiamati ad individuare le correlazioni tra i diversi item, esprimendo un giudizio sulla correlazione tra item diversi (ad esempio, un valutatore può considerare migliore qualcuno molto orientato al risultato ma che non rispetta le procedure piuttosto che chi fa poche pratiche ma rispetta le procedure ecc.).

L'idea alla base di questa proposta è la preferenza degli esperti a valutare in modo relativo piuttosto che in assoluto. Uno strumento che si potrebbe utilizzare al fine di implementare quest'intervento è un modello mutuato dall'SAHP (*Sparse Analytic Hierarchy Processes*, utilizzato in Computer Science), che permetta di comporre una gerarchia delle preferenze espresse dal valutatore in base alla comparazione degli item (valutati ciascuno sulla base degli altri) (Oliva et al., 2019).

Un secondo profilo meritevole di approfondimento riguarda le conseguenze derivanti dalla struttura a cascata del processo di valutazione.

1.2 Affiancamento della valutazione tra pari alla valutazione gerarchica

È possibile considerare nell'ottica di mitigazione del *leniency error* e dell'effetto alone che possono attivarsi nell'intero processo di valutazione, anche l'affiancamento al metodo di valutazione gerarchica quello della valutazione tra pari.

Questo strumento è particolarmente importante se si pensa al suo possibile ruolo anche all'interno di una strategia volta ad incidere sulla concezione della valutazione all'interno dell'organizzazione, cambiando l'atteggiamento sospettoso e negativo in una prospettiva in grado di apprezzarne l'opportunità e le potenzialità.

Nella prospettiva di implementazione di questa duplice modalità all'interno del sistema di valutazione ISPRA, vale la pena, però, sottolineare alcune criticità.

¹ Si veda in particolare la sezione 2.2 del presente report.

Tabella 1 - Dati relativi alla sottoscrizione delle valutazioni da parte dei valutati per gli anni 2019 e 2020

Sottoscritta?	2019		2020		Totale 19/20	
No	196	19.3%	154	14.8%	350	17.0%
Sì	435	42.8%	424	40.6%	859	41.7%
Non inviata	385	37.9%	465	44.6%	850	41.3%
Totale	1016	100.0%	1043	100.0%	2059	100.0%

La tabella qui sopra riporta i dati relativi alla sottoscrizione delle valutazioni da parte dei valutati per gli anni 2019 e 2020. Si nota come in entrambi i casi una porzione importante dei dipendenti (38% nel 2019, aumentata al 45% nel 2020) non solo non ha accettato la valutazione ricevuta, ma non ha proprio interagito con il sistema di valutazione nella sua interezza, segnalando un atteggiamento di rifiuto generalizzato verso il sistema di valutazione in quanto tale diffuso nell'organizzazione.

Ulteriore riprova di questo atteggiamento si può ritrovare nei commenti lasciati dai valutati che hanno apertamente rifiutato di sottoscrivere la valutazione ricevuta: molti commenti esprimono apertamente il rifiuto del sistema di valutazione delle performance indipendentemente dal voto ricevuto. In vari casi, questo rifiuto avviene sotto forma di azione collettiva, tramite l'inserimento di un commento standard identico da parte di vari dipendenti.

Questo atteggiamento generalizzato potrebbe rappresentare un ostacolo per eventuali tentativi di inserire da subito un sistema di valutazione tra pari (come già testimoniato dallo scarso successo della sperimentazione già effettuata in ISPRA) in un contesto caratterizzato da una cultura organizzativa ancora ostile al concetto di valutazione dei dipendenti. Riguarda dunque una prospettiva di cambiamento comportamentale a lungo termine, che richiede alcuni passi intermedi che possono essere delineati come riportato nei due seguenti paragrafi.

1.3 Aumento della consapevolezza con interventi di formazione mirati

L'attività formativa concernente la valutazione dovrà comprendere tutti i livelli dell'organizzazione attraverso moduli orientati alle singole categorie. Alla luce delle evidenze derivanti sia dalle valutazioni interne sia alle due survey si intende procedere alla organizzazione di percorsi formativi mirati.

Costruire dei laboratori formativi attraverso i quali misurare gli effetti del leniency error e dell'effetto alone, avviando così lo sviluppo di una consapevolezza che contrasti la tendenza all'uniformità nelle valutazioni (Sellier et al., 2019).

La proposta di intervenire con una specifica formazione è finalizzata, però, anche a moderare l'effetto di una possibile "valutazione a cascata", intesa come il fenomeno per cui un dirigente possa essere valutato dal OIV (o dal DG) anche in base alla valutazione (positiva o negativa) che il dirigente stesso ha operato sui suoi sottoposti.

Posto che sarà necessario prima di tutto indagare se e quando questo si verifica nella realtà (esistono, ad esempio, amministrazioni dove è sicuro che questa relazione non esista), appare evidente che, indipendentemente dal fatto che sia vero o falso, è più facile dare solo giudizi positivi perché è più economico (nel senso che richiede meno sforzo) ed è più gratificante (il dirigente che dà giudizi positivi è chiaramente più amato di chi dà anche giudizi negativi). In entrambi i casi si tratta di contenere, se possibile, il cosiddetto *self serving bias* (Wiekens & Stapel, 2010; Wang et al., 2017).

Per "quantificare" il precedente "se possibile", la nostra ipotesi è che maggiore è il senso di appartenenza all'ente, maggiore è la probabilità di contenere il *self serving bias*. L'idea è di approfondire con laboratori formativi gli aspetti sociali delle scelte individuali (un possibile strumento da utilizzare è quello del "training attraverso vignette", si veda a proposito Terry et al., 2017) così da attivare nei valutatori

l'identificazione con un livello organizzativo più ampio, predisponendoli ad abbandonare scelte opportunistiche a livello individuale o di network locale.

1.4 Modifiche preliminari al criterio di differenziazione di giudizio

Alla luce di quanto emerso dall'analisi delle valutazioni degli anni 2019 e 2020, si è pensato di intervenire modificando il criterio secondo il quale viene assegnato il valore del criterio di differenziazione di giudizio per il valutatore. Si è inoltre discussa la possibilità di fornire dei feedback ai valutatori sulle valutazioni effettuate negli anni precedenti, per evidenziare le opportunità di miglioramento. Viste le risorse limitate a disposizione, quest'ultima questione è per ora rimasta in sospeso.

Secondo il nuovo metodo di valutazione discusso, l'assegnazione del valore per il criterio di differenziazione di giudizio per il valutatore avviene tramite due componenti a cui viene attribuito peso identico, per un totale fino al 15%.

La decisione di introdurre il criterio di differenziazione dei giudizi attraverso la composizione di due componenti deriva da due necessità: da un lato, di migliorare il precedente metodo di valutazione del criterio di differenziazione che non risultava sufficientemente sensibile da scoraggiare l'attribuzione di valutazioni tendenti all'uniformità e che comunque consentiva al valutatore di evitare la penalizzazione tramite l'assegnazione strategica di poche valutazioni "forzate"; dall'altro, di essere in grado di individuare direttamente dai risultati della valutazione un 30% di dipendenti aventi diritto alla maggiorazione secondo la normativa contrattuale. Il nuovo metodo rende più esplicite tali aspettative nei confronti del processo di valutazione e punta ad incentivare un maggiore allineamento del comportamento dei valutatori a queste aspettative.

Componente I: Individuazione delle eccellenze

Se in seguito alla valutazione, è possibile individuare un 30% dei dipendenti da premiare dalla graduatoria senza creare ambiguità, il valutatore riceve un valore pari del 7,5% per la componente I, altrimenti riceve un valore pari a zero. Isolare questo criterio rende più evidente l'obiettivo di non avere situazioni di ex aequo in cui più dipendenti della percentuale prevista avrebbero diritto alla maggiorazione secondo la graduatoria, e permette di adattare facilmente tale percentuale del 30% ad eventuali cambiamenti contrattuali senza dover rivedere l'intero sistema di valutazione.

Componente II: Differenziazione

La differenziazione delle valutazioni viene premiata su una scala continua dallo 0 al 7,5% secondo la formula:

$$\% \text{ premio} = 7,5\% \times \left(1 - \frac{\sum \text{ripetizioni}_{\text{eccesso}_i}}{n_{\text{valutati}} - \frac{n_{\text{valutati}}}{n_{\text{voti}}}} \right)$$

dove:

$$\text{ripetizioni}_{\text{eccesso}_i} = \begin{cases} f_{\text{voto}_i} - \frac{n_{\text{valutati}}}{n_{\text{voti}}} & \text{se } f_{\text{voto}_i} - \frac{n_{\text{valutati}}}{n_{\text{voti}}} \geq 0 \\ 0 & \text{se } f_{\text{voto}_i} - \frac{n_{\text{valutati}}}{n_{\text{voti}}} < 0 \end{cases}$$

$$f_{\text{voto}_i} = \text{frequenza del voto } i$$

$$n_{\text{voti}} = \text{numero voti considerati "assegnabili" (almeno sufficienti)}$$

$$n_{\text{valutati}} = \text{numero dipendenti valutati dal dirigente con voto sufficiente}$$

La formula favorisce una distribuzione lineare delle valutazioni, in cui il valutatore fa uso di tutto il range di voti disponibile e tenta di concentrare il meno possibile le valutazioni su un numero ristretto di valori. Allo stesso tempo, la formula tiene anche conto del numero di dipendenti da valutare per il singolo valutatore, per non penalizzare chi è costretto a riutilizzare più volte gli stessi voti a causa dell'alto numero di dipendenti che si trova a dover valutare. Il calcolo di *nvalutati* esclude di proposito le insufficienze

(valutazione complessiva < 10) per evitare di creare incentivi all'assegnazione di insufficienze non necessarie. In questo modo, l'insufficienza rimane a disposizione come strumento per evidenziare criticità individuali (come già emerso dall'analisi delle valutazioni degli anni precedenti) il cui ricorso non inficia il risultato per il valutatore stesso, non determinando quindi la necessità di attribuire valutazioni "forzate" al solo fine di migliorare la propria prestazione.

Vista l'importanza attribuita alla maggiore esplicitazione delle aspettative dietro a questo cambiamento, presentare il nuovo metodo in una maniera il più efficace possibile risulta necessario al successo della modifica. Per questo motivo, testare varie formulazioni della presentazione del sistema in maniera sperimentale, nei prossimi passi da fare nel contesto della convenzione Unità di Scienze Comportamentali SNA e ISPRA, può essere utile al fine di individuare le strategie di presentazione più efficaci per gli obiettivi di cui sopra.

Conclusioni

L'analisi comportamentale ha contribuito ad evidenziare alcune debolezze del processo valutativo esaminando il rapporto tra organizzazione, valutatori e valutati.

Si sono presi in considerazione tre diversi livelli di analisi: 1) chi valuta, 2) come si valuta, 3) cosa si valuta.

Cercando di comprendere le relazioni tra i diversi livelli, si è deciso di mirare ad incentivare una valutazione del merito differenziata e disincentivare un modello di valutazione che esita nella distribuzione di incentivi uguali per tutti.

Per quanto riguarda la valutazione della performance individuale, uno degli aspetti innovativi del Sistema utilizzato in via sperimentale nel 2020 è il riflettore posto su quelli che sono definiti i “comportamenti organizzativi” richiesti per il più efficace svolgimento delle funzioni assegnate, richiamando l'attenzione non solo sui risultati, ma anche sul “come” questi risultati si raggiungono. L'enfasi sugli aspetti comportamentali viene in particolare sottolineata dalla scelta del metodo plurifattoriale per la valutazione della prestazione, fondato, nello specifico, sulle competenze professionali dimostrate (comportamenti organizzativi) e la performance organizzativa dell'unità di appartenenza. Considerando che i comportamenti possono essere interpretati come prodotti dalle competenze agite dalla persona, diventano il segnalatore della disponibilità e dell'uso delle competenze stesse. In questa visione del modello teorico misto di valutazione, il concetto di performance diventa più ampio e abbraccia, insieme ai risultati/obiettivi (più o meno quantificabili), i comportamenti messi in atto per realizzarli.

Tra gli elementi chiave di processo e di contenuto utilizzati per identificare la tipologia di valutazione da utilizzare, infatti, due punti in particolare delle Linee guida del Dipartimento della funzione pubblica vanno in questa direzione: (1) l'enfasi sui processi organizzativi e non solo sui risultati, in modo tale da poter valutare gli *outcome* (impatti) anziché gli output (risultati), poiché la producibilità degli impatti attesi di lungo termine è possibile proprio grazie al controllo dei processi organizzativi, dei comportamenti e delle qualità delle scelte effettuate; (2) il coinvolgimento delle risorse umane, con l'intento di cogliere, attraverso l'implementazione del modello, un'occasione di formazione individuale e di creazione di nuove convinzioni e valori manageriali nei gruppi.

Affinché l'implementazione del Sistema possa essere perseguita in modo efficace, è stato previsto, sin dalla prima sperimentazione, un accompagnamento graduale, finalizzato alla gestione del cambiamento e ad evitare crisi di rigetto. La collaborazione con l'Unità di Scienze Comportamentali della SNA si inserisce, di fatto, in questo percorso, con il fine di identificare aspetti comportamentali strategici e progettare interventi che consentano di correggere alcune distorsioni cognitive che possono caratterizzare il processo di valutazione, rendendo consapevoli gli attori dei bias a cui possono essere soggetti e contribuendo all'azione di contrasto alla manipolazione delle attività di valutazione a fini esogeni rispetto l'organizzazione e l'obiettivo di miglioramento dei processi e degli *outcome* organizzativi (ingerenze esterne, strategie autoreferenziali, etc.).

In particolare, ci si è concentrati sui valutatori e si è promosso il coinvolgimento degli individui chiamati a valutare anche in fase di implementazione e monitoraggio del Sistema, permettendo e sollecitando confronti sulle diverse aspettative e sulle prospettive riguardo cosa valutare, per avviare un processo di architettura concertata delle scelte relative alla performance e agli item da valutare. In altri termini, si è avviato un processo che mira ad esplicitare le aspettative degli attori coinvolti nella valutazione, permettendo allo stesso tempo di determinare in che misura le performance individuali sono autonome ed in che misura dipendono da aspetti organizzativi, considerando che tale dipendenza può avere caratteristiche non solo motivazionali, ma anche strutturali.

Nello specifico, le modifiche sia agli strumenti utilizzati sia all'intero processo di valutazione, così come il percorso formativo riguardante la valutazione di performance, sono state proposte integrando attività di ricerca ed attività di formazione ed accompagnando l'amministrazione nell'analisi della

domanda formativa e nelle risposte alle esigenze di cambiamento emerse. Si tratta però di un processo che necessita di essere proseguito per allineare gli obiettivi dell'organizzazione con quelli dei valutatori grazie alla possibilità di far emergere le divergenze di credenze esistenti in modo trasparente e consapevole. In altri termini, l'idea è di proseguire con l'approccio comportamentale, affiancando le diverse strategie di intervento più strutturale che sono state individuata, per fare in modo che la riflessione su questi aspetti diventi momento di formazione e spinta verso possibili cambiamenti culturali. L'obiettivo è di progettare interventi di *norm-nudging*, a partire dalla misurazione e dall'analisi delle attuali aspettative sociali, misurazione e analisi che devono necessariamente proseguire in modo concertato, attraverso focus group formativi e laboratori sistematici.

Un ulteriore sviluppo della ricerca riguarda l'approfondimento della prospettiva dei valutati, fondamentale, al pari di quella dei valutatori sui quali ci si è focalizzati, per il cambiamento culturale necessario ai fini dell'accettazione dell'intero Sistema. A partire da una distinzione tra processo decisionale del valutatore e processo decisionale del valutato, si mirerà, cioè, ad utilizzare l'approccio comportamentale per indagare empiricamente gli aspetti motivazionali che intervengono nel relazionarsi, da parte dei singoli individui all'interno dell'organizzazione, con la valutazione, le sue potenzialità e le sue conseguenze. Grazie a questa nuova prospettiva sarà inoltre possibile investigare come l'interazione di una pluralità di componenti cognitive, relazionali e di contesto spinga l'individuo a mettere in atto un determinato comportamento non solo nei confronti del Sistema di valutazione, ma anche del proprio ruolo nello svolgimento delle attività lavorative.

Il fine è quello di integrare la prospettiva sui valutatori e quella sui valutati, per progettare e implementare strategie comportamentali di supporto ad una *governance* collaborativa, che implichi la condivisione degli obiettivi nella fase della programmazione (i.e. definizione e comunicazione delle aspettative), e la partecipazione alla definizione del Sistema di valutazione, in un percorso di apprendimento continuo (i.e. fasi successive e di feedback).

Riferimenti bibliografici

S.C. Andersen., & M.H. Larsen, (2016). *Cognitive Biases in Performance Evaluations*. Journal of Public Administration Research and Theory, 26(4), 647- 662.

S.C. Andersen. & M. Hjortskov (2016). *Cognitive Biases in Performance Evaluations*. Journal of Public Administration Research And Theory; Vol. 26 No. 4, 647–662.

C.H. Antoni, X. Baeten, S.J. Perkins, J.D. Shaw, M. Vartiainen (2017). Reward Management: *Linking Employee Motivation and Organizational Performance*. Journal of Personnel Psychology, 16(2), 57–60.

P.P. Baard, E.L. Deci, R.M. Ryan, (2004). *Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings*, Journal of Applied Social Psychology, 34(10), 2045–2068.

M. Barbieri, G. Girosante and G. Valotti (2017). *Comparative Analysis of the Performance Evaluation Systems of Public Sector Employees*, Egea Edizioni.

R.P. Battaglio, P. Belardinelli, N. Belle, P. Cantarelli (2018). *Behavioral Public Administration ad fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations*. Public Administration Review, Vol. 79, Iss. 3, 304–320.

P. Belardinelli, N. Belle, M. Sicilia, I. Steccolini (2018). *Framing Effects under Different Uses of Performance Information: An Experimental Study on Public Managers*. Public Administration Review, Vol. 78, Iss. 6, 841–851.

N. Belle, P. Cantarelli, P. Belardinelli (2017). *Cognitive Biases in Performance Appraisal: Experimental Evidence on Anchoring and Halo Effects With Public Sector Managers and Employees*. Review of Public Personnel Administration, Vol. 37(3), 275–294.

E.M. Berger, C. Hrabring, D. Sliwka (2013). *Performance Appraisals and the impact of forced distribution – An experimental investigation*. Management Science, 59(1), 54-68.

C. Bicchieri (2017). *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*, Oxford University Press.

J.S. Bowman, (1994). *At last, an alternative to performance appraisal: Total quality management*. Public Administration Review, 54, 129-136. J.S. Bowman (1999). *Performance appraisal: Verisimilitude trumps veracity*. Public Personnel Management, 28, 557-576.

R.W. Brown (1982). *Performance appraisal: A policy implementation analysis*. Review of Public Personnel Administration, 2, 69-85.

B. Bruno, M. Faggini, A. Parziale (2017). *Motivation, Incentives and Performance: An Interdisciplinary Review*. International Journal of Business and Management; Vol. 12, No. 12.

F. Cafaggi (2020) *Gerarchia e processi di valutazione della performance: la peer review come strumento di debiasing*, unpublished paper on file with the author

F. Cafaggi (2021) *Processi di valutazione e organizzazione amministrative: aspettative e performance organizzativa*, unpublished paper on file with the author

C. Castelfranchi & E. Lorini (2003). *Cognitive Anatomy and Functions of Expectations*. In F. Schmalhofer, R. Young, G. Katz (Eds.) Proceedings of Euro CogSci 03, Routledge.

P. Cantarelli, N. Belle, P. Belardinelli (2020). Review of Public Personnel Administration, Vol. 40(1), 56–81.

O. Chinotti, A. Festa, *Nuovi strumenti per lo sviluppo delle persone nelle organizzazioni: la tecnica della valutazione a 360 °*. Milano: Franco Angeli, 2008.

M. Coccia (2019). *Intrinsic and extrinsic incentives to support motivation and performance of public organizations*. Journal of Economics Bibliography, Vol. 6, Issue 1.

M. Cohen & C. Ranganath (2007). *Reinforcement learning signals predict future decisions*. The Journal of neuroscience, 27(2), 371-378.

Dipartimento della Funzione Pubblica (2017). *Linee guida per il Sistema di misurazione e valutazione della performance dei Ministeri*, n. 2/2017.

Dipartimento della Funzione Pubblica (2019). Linee guida per la misurazione e valutazione della performance individuale, n.5/2019.

European Commission. Directorate-General for Employment, Social Affairs and Inclusion (2015). Quality of Public Administration. Toolbox for Practitioners.

A. Furnham & H.C. Boo (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40, 35-42.

G.T. Gabris & D.M. Ihrke (2000). *Improving employee acceptance toward performance appraisal and merit pay systems: The role of leadership credibility*. Review of Public Personnel Administration, 20, 41-53.

D. Kahneman (2011). *Thinking, fast and slow*. New York City, NY: Macmillan.

J.S. Kane, S.H. Bernardin, P. Villanova, J. Peyrefitte (1995). *Stability Of Rater Leniency: Three Studies*, Academy of Management Journal, Vol. 38, No. 4, s.1039.

R.P. Larrick (2004). Debiasing. In Rachel A. Powsner & Edward R. Powsner (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316-338).

D.C. Martin & K.M. Bartol (1986). *Training the raters: A key to effective performance appraisal*. Public Personnel Management, 15, 101-109.

L. Marengo, C. Pasquali (2010). *How to get what you want when you do not know what you want: A model of incentives, organizational structure and learning*. LEM Working Paper Series, No. 2010/08.

Ng, K. Y., Koh, C., Ang, S., Kennedy, J. C., Chan, K. Y. (2011). *Rating leniency and halo in multisource feedback ratings: testing cultural assumptions of power distance and individualism-collectivism*. J. Appl. Psychol. 96, 1033.

J.P. O' Doherty (2004). *Reward representations and reward-related learning in the human brain: insights from neuroimaging*. Current Opinion in Neurobiology, 14, 769-776.

OECD (2017). *Engaging Public Employees for a High-Performing Civil Service*, Organization for Economic.

G. Oliva, R. Setola, A. Scala (2019). *Sparse analytic hierarchy process: an experimental analysis*. Soft Comput. 23, 2887-2898.

G. Ostinelli (2005). *Motivazione e comportamento: le variabili psicologiche necessarie per raggiungere obiettivi*. Trento: Erickson.

J.B. Rowe, D. Eckstein, T. Braver, A.M. Owen (2008). *How Does Reward Expectation Influence Cognition in the Human Brain?* Journal of Cognitive Neuroscience, 20(11).

J. So, A. Chethana, H. DaHee, A. Nidhi, D. Adam, M. Durairaj (2015). *The Psychology of Appraisal: Specific Emotions and Decision-Making*. Journal of Consumer Psychology. 25. 10.1016.

J.B. Soll, K.L. Milkman, J.W. Payne (2014) *A User's Guide to Debiasing*, In Gideon Keren and George Wu (Eds.) Handbook of Judgment and Decision Making, Wiley-Blackwell.

R.H. Thaler, C.R. Sunstein (2018). *La spinta gentile. La nuova strategia per migliorare le nostre decisioni su denaro, salute, felicità*, Feltrinelli, Milano.

T.J. Thorsteinson, J. Breier, A. Atwell, C. Hamilton & M. Privette (2008). *Anchoring effects on performance judgments*. Organizational Behavior and Human Decision Processes, 107, 29-40.

J. Taylor & J. H. Westover (2011). *Job satisfaction in the public service: The effects of public service motivation, workplace attributes and work relations*. Public Management Review, 13, 731-751.

W. Van Eerde & H. Thierry (1996). *Vroom's expectancy models and work-related criteria: A meta-analysis*. Journal of Applied Psychology, 81(5), 575-586.

R. Viale (2018). *Oltre il nudge. Libertà di scelta, felicità e comportamento*, Il Mulino, Bologna.

Vroom, V.H. 1964. *Work and motivation*. New York: Wiley.

G. Yildiz, A. Baltaci (2009). *Leniency and Severity Errors in Performance Appraisal in the Context of Collectivist and Individualist Culture*, International Symposium on Sustainable Development, Saraybosna, Bosna-Hersek, pp. 342-348.



SNA

Presidenza del Consiglio dei Ministri
Scuola Nazionale dell'Amministrazione