



UNIVERSITÀ DEGLI STUDI DI PADOVA

**Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia Applicata
(FISPPA)**

**Corso di laurea Magistrale in Psicologia Sociale, del Lavoro e della
Comunicazione**

Tesi di laurea Magistrale

***Invarianza cross-culturale della chiamata professionale:
uno studio empirico.***

Cross-cultural invariance of career calling: an empirical study.

Relatore

Prof. Michelangelo Vianello

Laureanda: Melania Maiorelli

Matricola: 2014768

Anno Accademico 2022-2023

Sommario

1.	IL COSTRUTTO DI CHIAMATA PROFESSIONALE	6
1.1	Sviluppo, diffusione e antecedenti del costrutto di chiamata	6
1.1.1	L'approccio neoclassico alla chiamata.....	7
1.1.2	L'approccio moderno alla chiamata	9
1.1.3	Presupposti del costrutto di chiamata.....	10
1.1.4	Differenze tra il costrutto di chiamata e costrutti affini	12
1.2	Le diverse teorie di chiamata professionale	13
1.2.1	<i>Work as Calling Theory</i> (WCT).....	16
1.3	Le diverse misure di chiamata professionale	20
1.3.1	<i>Unified Multidimensional Calling Scale</i> (UMCS)	22
2.	L'INVARIANZA DELLA MISURA	25
2.1	I problemi di comparabilità nella ricerca cross-culturale	25
2.1.1	Il costrutto di <i>bias</i> nella ricerca cross-culturale.....	26
2.1.2	Il costrutto di equivalenza nella ricerca cross-culturale	30
2.2	Il costrutto di misura secondo la Teoria Classica dei Test	32
2.3	Il costrutto di invarianza della misura	36
2.3.1	L'invarianza della misura configurale, metrica, scalare e stretta	38
2.4	L'Analisi Fattoriale Confermativa (CFA).....	43
2.4.1	L'Analisi Fattoriale Confermativa di primo e di secondo ordine.....	46
2.4.2	Gli indici di <i>fit</i> nell'Analisi Fattoriale Confermativa	48
3.	STUDIO EMPIRICO	53
3.1	Obiettivi e Ipotesi di ricerca	53
3.2	Metodo	53
3.2.1	L'approccio Open Science	53
3.2.2	Partecipanti.....	55

3.2.3 Procedura.....	62
3.2.4 Misure e Strumenti	62
3.3 Analisi dei dati	65
3.4 Risultati	65
3.4.1 Risultati dell'analisi fattoriale confermativa di secondo ordine	65
3.4.2 Risultati dell'analisi dell'invarianza della misura multi-gruppo	68
3.5 Discussione	74
Bibliografia	79

1. IL COSTRUTTO DI CHIAMATA PROFESSIONALE

1.1 Sviluppo, diffusione e antecedenti del costrutto di chiamata

Il concetto di lavoro ha assunto, nel corso della storia, diversi ruoli. Le norme legate all'epoca e alla cultura di appartenenza aiutano infatti a fissare aspettative sociali, influenzando il significato soggettivo che il lavoro ha per gli individui e regolando le interazioni sociali (Ruiz Quintanilla & Wilpert, 1988).

Nel corso dell'Antica Grecia il lavoro era considerato degradante e riservato agli schiavi: esso era una maledizione che impediva all'umanità di impegnarsi in attività più sublimi e meritevoli della mente e dello spirito. Questa visione continuò a dominare gli insegnamenti filosofici e religiosi per tutto il Medioevo, mentre nel Rinascimento il lavoro iniziò ad essere concettualizzato come mezzo con cui l'umanità poteva paragonarsi al divino. (Dik & Duffy 2009).

Il costrutto di *calling* e l'idea che il lavoro possa essere concettualizzato come una chiamata o una vocazione trova le sue origini nel Sedicesimo secolo. Fino a quel momento, la chiamata vocazionale era riservata alla sfera clericale, ed indicava la chiamata del singolo a prendere parte al ministero della chiesa o l'universale chiamata delle scritture (Weber, 1930). Il Riformatore Protestante Martin Lutero estese ed ampliò questa definizione di chiamata, utilizzando il termine per riferirsi a qualsiasi tipo di lavoro produttivo. Secondo Lutero, la fedele esecuzione di tale lavoro avrebbe portato a benefici sia agli occhi di Dio, sia per il resto della comunità (Dik & Duffy, 2009). Il concetto di chiamata introdotto da Lutero trasforma così il lavoro da “male necessario” a “offerta divina”. Il sociologo e filosofo Max Weber, analizzando gli scritti di Lutero nella sua

opera del 1904 “L’etica protestante e lo spirito del capitalismo” (Weber, 1930), nota come il Riformatore Protestante, nel tradurre in tedesco la Bibbia, utilizza la parola “chiamata” (*beruf* nella versione originale tedesca) per indicare lo scopo preordinato di Dio per ogni individuo. Questa parola unisce la nozione secolare di lavoro e della propria occupazione con la nozione religiosa di scopo affidato da Dio ad ogni individuo. In questo modo, Lutero crea un concetto completamente nuovo: il fedele lavoro secolare o laico rappresenta il più alto bene morale possibile per l’uomo. Secondo questa concezione, Dio vuole che l’essere umano lavori duramente alla propria chiamata (Bunderson & Thompson, 2009).

L’idea viene successivamente ripresa dal Protestante Calvino, che elevò ulteriormente il concetto di lavoro, introducendo l’individualità della chiamata. Per Lutero, il tipo di chiamata dipendeva solamente dallo stadio in cui la persona si trovava nella vita (Weber, 1930), mentre per Calvino essa dipende dagli specifici e particolari talenti, donati alla persona da Dio. La chiamata diventa un modo per sfruttare il proprio talento per il bene della comunità e dell’Umanità, e diventa quindi un dovere del singolo quello di trovare e sviluppare la propria vocazione. (Dik & Duffy, 2009). Questa origine della concettualizzazione di chiamata influenzerà l’approccio neoclassico al costruito, in opposizione all’approccio moderno di *calling*. Di seguito, i due approcci spiegati e a confronto.

1.1.1 L’approccio neoclassico alla chiamata

L’approccio neoclassico alla chiamata è ancorato alla nascita del concetto di chiamata che trae le sue origini dalla Rivoluzione Protestante nell’Europa Cristiana, che chiameremo “approccio classico”. Thompson e Bunderson (2019) appuntano tre

principali elementi specifici dell'approccio classico. In primis, la visione classica presume che ogni individuo sia destinato ad un certo lavoro. Secondariamente, spetta all'individuo individuare e accogliere tale chiamata, la quale è un dovere che deriva dall'obbligo di glorificare Dio ed essere proficuamente coinvolto nel servizio alla propria famiglia e comunità. Infine, la visione classica implica che la ricerca della propria chiamata sia un processo che inizia con la valutazione dei propri talenti, passioni e opportunità di vita (Thompson & Bunderson, 2019). L'approccio neoclassico mantiene queste premesse, riformulandole in una concezione secolare o laica, che non presume o richiede un'esplicita credenza nel divino (Thompson & Bunderson, 2019). Rimangono immutati gli elementi chiave della chiamata classica quali il destino, il dovere prosociale e il concetto di scoperta della *calling* (Bunderson & Thompson, 2009). L'approccio neoclassico è tutt'ora utilizzato dai ricercatori: per Bellah et. al (1985) ogni chiamata legittima è percepita come un contributo per il bene comune, mentre per Bunderson & Thompson (2009) alcuni percorsi occupazionali risultano destinati in quanto essi sono perfettamente aderenti agli interessi e capacità del singolo, nonché gli permettono di contribuire nella società. Dik & Duffy (2009, p. 427) affermano che la chiamata è un richiamo trascendente, originata oltre e al di fuori del sé, che motiva attraverso valori e obiettivi orientati all'altro (Thompson & Bunderson, 2019). La chiamata neoclassica si focalizza quindi su fattori esterni all'individuo rispetto al concetto di vocazione che non include la trascendenza ma solo la prosocialità e lo scopo nella vita di ognuno (Dik & Duffy, 2009).

1.1.2 L'approccio moderno alla chiamata

In contrasto all'approccio neoclassico alla chiamata, l'approccio moderno è più allineato ai valori e alle sensibilità moderne, staccandosi dalle origini classiche del costrutto (Thompson & Bunderson, 2019). Le concettualizzazioni moderne di chiamata riflettono un'enfasi sull'autoespressione e l'autorealizzazione, in linea con la moderna enfasi sull'individualismo. Secondo questa visione, viene meno il senso di dovere o obbligo nei confronti della società: la chiamata risulta quindi essere unicamente un'espressione di passioni e interessi personali, e risulta acquisire importanza dal momento in cui è significativa per l'individuo e non per la società e il suo beneficio (Thompson & Bunderson, 2019). La chiamata perde inoltre il concetto di destino o necessità di scoperta del proprio talento e abilità che aveva acquisito nella concezione neoclassica, e acquisisce la caratteristica di espressione personale del sé. Essa è scelta e adottata dal singolo per la propria aspirazione (Thompson & Bunderson, 2019). Questo tipo di chiamata risulta quindi incentrata su bisogni e desideri personali (Dobrow & Tosti-Kharas, 2011). Sebbene la concezione di chiamata moderna si distacchi dalle radici Protestanti di origine, anche in essa sono evidenti le radici Cristiane Occidentali, soprattutto nelle sue componenti chiave di pervasività, identità, sacrificio e passione (Vianello et al., 2022).

Le due concettualizzazioni di chiamata neoclassica e moderna sono da considerarsi come due estremi di un continuum (Dik & Shimizu, 2019), e il dibattito sopra quale definizione sia quella "corretta" è più un dibattito linguistico che scientifico: Hirshi (2011) ha dimostrato che anche i partecipanti alla ricerca sono divisi in termini di significato attribuito al costrutto. Il fatto che la definizione di *calling* sia soggetta ad un

continuo dibattito è da attribuirsi alla natura dinamica e in continua evoluzione del costrutto (Wrzesniewski, 2012). Thompson & Bunderson (2019) affermano inoltre che come quello che inizialmente è nato come “dibattito” tra la concezione neoclassica e quella moderna di chiamata dovrebbe ad oggi essere considerata un’integrazione tra le due prospettive: secondo la maggior parte degli studiosi della chiamata, considerare solamente una delle due prospettive senza tenere in considerazione l’altra risulterebbe una perdita di importanti dimensioni del costrutto (Thompson & Bunderson, 2019). Come affermato da Abraham Maslow (1967), la convinzione che un lavoro sia più dei costrutti di *job* e *career* può risultare non solo da un senso di necessità interna (come affermato dalla visione moderna) o necessità esterna (come affermato dalla visione neoclassica), ma da una narrativa personale che unisce e sintetizza le due visioni (Thompson & Bunderson, 2019).

1.1.3 Presupposti del costrutto di chiamata

Per rendere più chiaro il costrutto di chiamata, si riportano i cinque presupposti impliciti per la concettualizzazione di *calling*, che Dik e Duffy (2009) rendono espliciti. Il primo assunto riguarda la concezione degli individui come agenti attivi, capaci di intenzionalità, autoregolazione, lungimiranza e autoriflessione (Bandura, 2006). L’agency è inoltre influenzata da una complessa interazione di fattori biologici, ambientali e spirituali (Bandura, 1982). Successivamente, si assume che gli esseri umani siano organismi creatori di significato (Baumeister, 1991), costruttori di significati globali e particolari per le proprie esperienze di vita. Tale assunzione di significato è fatta in modo sia esplicito che implicito. Come terzo presupposto, si assume che gli individui assumano numerosi ruoli sociali (Super, 1980), che interagiscono tra loro in modo

complesso. Il lavoro occupa uno di questi ruoli sociali. Si assume che gli esseri umani, per necessità, vivano in società legate da bisogni comuni e che le attività legate al lavoro abbiano implicazioni sociali dirette o indirette, di varia intensità (Blustein, 2006; Hardy, 1990). Infine, si assume che gli ostacoli alla ricerca di significato e di uno scopo sul lavoro siano presenti su più livelli (ad esempio, individuali, organizzativi, di società), ma che essi siano suscettibili al cambiamento (Dik & Duffy, 2009).

Gli autori (Dik & Duffy, 2009) esplicitano inoltre alcune implicazioni fondamentali per il costrutto di chiamata. Innanzitutto, la *calling* non è qualcosa che l'individuo scopre una volta per tutte, ma coinvolge un processo continuo di valutazione del significato e degli scopi delle attività svolte, nell'ottica del possibile contributo per la società e per il bene comune. La *calling* non si limita quindi solo alla scelta della carriera, e quindi al comportamento decisionale iniziale, ma include anche le scelte che l'individuo fa all'interno della carriera stessa, e come questa possa essere perseguita in modo tale da connettere le attività eseguite sul lavoro al significato e gli scopi orientati all'altro e al bene comune (Dik & Duffy, 2009). Questo rende il costrutto di chiamata rilevante anche per coloro che non abbiano perseguito la propria carriera ideale: attraverso il *job crafting*, questi individui saranno in grado di riformulare e ridefinire le proprie attività e responsabilità lavorative in un modo che trasformi il proprio lavoro in una chiamata. (Duffy et al., 2018). I presupposti elencati portano alla seguente affermazione: ogni persona può potenzialmente avere una chiamata, e la chiamata può verificarsi in qualsiasi ambito lavorativo, anche se l'individuo non risulta direttamente coinvolto nel perseguire il benessere degli altri (Dik & Duffy, 2009). Wrzesniewski e Dutton (2001) avvalorano la precedente affermazione, dimostrando che gruppi di assistenti amministrativi, operatori delle pulizie ospedaliere, parrucchieri e dipendenti della ristorazione affermano di

avvicinarsi al proprio lavoro come una chiamata, o di aver messo in atto *job crafting* in modo da permettere loro di sperimentare un significato e scopo sul lavoro, per perseguire il benessere altrui (Wrzesniewski & Dutton, 2001; Wrzesniewski et al., 1997).

1.1.4 Differenze tra il costrutto di chiamata e costrutti affini

La chiamata presenta affinità con alcuni costrutti, dai quali è importante porre una differenziazione. I costrutti di *work centrality* (Dubin, 1956), *work commitment* (Loscocco, 1989), *career salience* (Greenhaus, 1971) e *work e job involvement* (Kanugo, 1982) fanno riferimento a vari aspetti dell'importanza psicologica che un individuo attribuisce alla propria carriera. Questi costrutti non catturano però la costruzione di senso e gli aspetti prosociali esaminati dalla *calling* (Dik & Duffy, 2009). Inoltre, il costrutto di *calling* si focalizza sui contributi prosociali raggiungibili attraverso il lavoro svolto e sull'identificazione di una forza interna o esterna che guidi gli individui verso una particolare carriera (Duffy et al., 2018). Questo permette di distinguere il costrutto dagli affini concetti di *meaningful work*, motivazione intrinseca e *sense of purpose at work* (Brown & Lent, 2016; Dik & Duffy, 2009). La chiamata si distingue inoltre dal costrutto di *personal engagement* sul lavoro (Kahn, 1990). Quest'ultimo rappresenta un coinvolgimento più episodico sul posto di lavoro, al contrario della natura stabile della *calling*. Infine, la chiamata si distingue dal costrutto di *workplace spirituality* (Giacalone & Jurkiewicz, 2003): questo fenomeno è definito come l'insieme dei valori presenti nella cultura aziendale che promuove l'esperienza di trascendenza dei dipendenti attraverso i processi di lavoro, facilitando il loro senso di connessione agli altri, in un modo che fornisca sentimenti di gioia e completezza (Dik & Duffy, 2009). È infine rilevante la distinzione tra il costrutto di chiamata e quello di vocazione. Il costrutto di vocazione è

storicamente molto legato a quello di chiamata: i termini, dapprima utilizzati interscambiabilmente, iniziano ad evolversi e ad essere distinti (Schuurman, 2004; Serow, 1994). Dik & Duff (2009), definiscono la vocazione come un approccio ad un particolare ruolo nella vita che sia orientato alla dimostrazione o alla derivazione di un senso di scopo o significato, e che presenta come fonte di motivazione valori e obiettivi orientati all'altro (Dik & Duffy, 2009). A partire da questa definizione possiamo affermare che, sebbene sia gli individui con chiamate sia quelli con vocazioni connettano il proprio lavoro ad un senso di scopo e ricerca di significato verso l'altro, solo gli individui con *callings* percepiscono l'impulso ad avvicinarsi al lavoro in questo modo come proveniente da una fonte esterna al sé (Dik & Duffy, 2009).

1.2 Le diverse teorie di chiamata professionale

Il costrutto di chiamata professionale, o *career calling*, è uno dei più consolidati e antichi nello studio del lavoro e delle organizzazioni. In letteratura, si possono osservare oltre 200 studi pubblicati nell'ultimo decennio (Duffy et al., 2018), con un interesse per la ricerca che ha preso piede tra il 2005 e il 2009 e un'esponentiale crescita per il costrutto di *work as a calling* dopo l'anno 2011 (Thompson & Bunderson, 2019). In una *literature review* pubblicata nel 2013, Duffy e Dik osservano una nuova e crescente attenzione per il costrutto di chiamata tra i ricercatori, consolidandone i temi e osservando i possibili futuri sviluppi. L'interesse per lo studio della chiamata professionale e della percezione di *work as a calling* matura con la crescita della ricerca nel campo, a seguito della dimostrazione dell'influenza del costrutto sull'aumento della percezione di significato e della soddisfazione negli ambiti della carriera, del lavoro e della vita tra numerosi

campioni di studenti universitari e adulti lavoratori (Duffy & Dik, 2013; Duffy et al., 2018). La letteratura sul tema nasce e si sviluppa in ambito multidisciplinare nei campi della psicologia del lavoro e delle organizzazioni, vocazionale, multiculturale e nel counseling (Duffy et al., 2018).

Bellah et al (1986) iniziano a definire il costrutto, articolando la distinzione tra tre orientamenti al lavoro: *job*, *career* e *calling*. I ricercatori definiscono *job* il lavoro utilizzato come mezzo per ottenere beni materiali e *career* il lavoro come mezzo per ottenere e aumentare potere e prestigio. Infine, Bellah et al (1986) definiscono *calling* la soddisfazione che si trae dal lavoro intrinsecamente appagante e dotato di valore sociale. Wrzesniewski et al. (1997) operazionalizzano i tre orientamenti al lavoro proposti e dimostrano come essi si distribuiscano in modo equo in un campione di dipendenti universitari: i ricercatori osservano come il campione che percepisce la *calling* risulti essere quello con i livelli più alti di *job satisfaction* e *life satisfaction* (Wrzesniewski et al., 1997).

Successivamente, molti ricercatori si sono approcciati ad una definizione di chiamata come costrutto a sé, e non in opposizione ad altri orientamenti al lavoro. Tra le definizioni che utilizzano la prospettiva moderna di chiamata, Dobrow e Tosti-Kharas (2011, p. 1005) definiscono la *calling* come “una passione profonda e significativa che le persone sperimentano verso un dominio”: questa definizione pone la chiamata come centrale per l’identità personale, promuovendo un impegno significativo verso i domini lavorativi che beneficino l’individuo, la propria famiglia o la società (Duffy et al., 2018). Praskova et al (2015, p. 93) inquadrano il costrutto di chiamata come “obiettivo di carriera di ordine superiore, perlopiù autoimposto, che genera significato e scopi per l’individuo”.

Hall e Chandler (2005, p. 160) definiscono la *calling* più generalmente come “cosa la persona vede come il proprio scopo nella vita”. Queste definizioni dimostrano come i ricercatori con approccio moderno alla chiamata si siano distanziati dalle radici religiose del termine, riformulando il costrutto come un approccio al lavoro altamente personale, significativo e intrinsecamente motivato (Dik & Shimizu, 2019). I ricercatori che abbracciano l’approccio moderno utilizzano definizioni perlopiù unidimensionali, basate sulla componente del significato e scopo del lavoro (Dik & Shimizu, 2019).

All’altro estremo del continuum, tra i ricercatori che adottano una prospettiva neoclassica di chiamata, troviamo la concettualizzazione di Dik & Duffy (2009, p. 427) che definiscono la *calling* “un richiamo trascendente, originato oltre e al di fuori del sé, per avvicinarsi a un particolare ruolo della vita in un modo orientato verso la dimostrazione o derivazione di un senso di scopo o significato, che detiene valori e obiettivi orientati verso gli altri come fonti primarie di motivazione”. Fornendo un’ulteriore definizione di stampo neoclassico, Bunderson & Thompson (2009, p. 38), definiscono il costrutto “il posto nella visione del lavoro della società che un individuo si sente destinato a ricoprire per un particolare dono o talento”. Queste concettualizzazioni di *calling* sono multidimensionali (Dik et. al, 2012), e considerano sia la componente del significato e scopo del lavoro, sia la componente legata ad una fonte trascendente di chiamata, un *caller*.

Al centro del continuum tra prospettiva neoclassica e moderna si collocano alcuni ricercatori che concettualizzano la chiamata come multidimensionale, considerando la componente del significato e scopo del lavoro e la componente di motivazione prosociale. Per questi *scholars*, così come per i moderni, non è necessario una fonte esterna di

chiamata. Ne sono un esempio la ricerca di Elangovan et al. (2010, p. 430), secondo cui la calling è definita come “modo di agire nella ricerca di intenzioni prosociali che incarnano la convergenza tra cosa un individuo vorrebbe fare, cosa dovrebbe fare e cosa realmente fa”, e la ricerca di Wrzesniewski et al (1997, p. 21) secondo cui essa è “il focus dell’individuo sulla soddisfazione per un lavoro gratificante e utile socialmente”.

Allontanandosi dalla divisione tra prospettiva neoclassica e moderna e approcciandosi al costrutto di calling da un punto di vista culturale, la ricerca dimostra come alcune concettualizzazioni e misure di calling più siano adatte a certi gruppi culturali rispetto che ad altri (Duffy et al., 2018). Ad esempio, gruppi di individui che stimano e attribuiscono importanza ad un credo religioso o ad aspetti spirituali possono essere attratti dall’aspetto di richiamo trascendente del costrutto rispetto ad altri gruppi (Dik, Duffy & Tix, 2012). Allo stesso modo, gruppi di individui che attribuiscono importanza all’individualismo possono essere attratti dalle dimensioni di significato e scopo più di altri gruppi, mentre gruppi che valorizzano il collettivismo possono essere più attratti dalla dimensione prosociale (i.e. servizio alla famiglia, alla comunità o alla società in generale) (Dik & Duffy, 2009).

1.2.1 *Work as Calling Theory* (WCT)

Nonostante il crescente interesse per la chiamata professionale, la conoscenza scientifica accumulata risulta limitata dall’assenza di un comune quadro teorico (Vianello et al. 2020). Se da una parte queste differenze concettuali sono descrivibili come un segno della natura dinamica e in costante sviluppo della ricerca sulla chiamata, dall’altra creano un ostacolo per lo sviluppo della teoria: diverse concettualizzazioni di chiamata hanno infatti portato allo sviluppo di diverse misure del costrutto (Duffy et al., 2018). Per questo

motivo, non risulta possibile stabilire la generalizzabilità dei risultati ottenuti dalla ricerca sviluppata utilizzando i numerosi modelli proposti (Vianello et al., 2020), e risulta necessario il tentativo di proporre un unico modello teorico di riferimento. A questo scopo, i ricercatori Duffy, Dik et al. pubblicano nel 2018 l'articolo "Work as a Calling: A Theoretical Model" sul Journal of Counseling Psychology (Duffy, Dik, Velez & Douglass, 2018). La ricerca nasce con l'intento di sintetizzare un modello teorico, *Work as Calling Theory* (WCT), e presenta (a) i predittori del vivere una chiamata professionale, (b) le variabili che moderano e mediano la relazione tra il percepire e il vivere una chiamata e (c) gli *outcome* positivi e i possibili *outcome* negativi (il cosiddetto lato oscuro della chiamata o *dark side of calling*), risultanti dal vivere la chiamata (Duffy et al., 2018). Di seguito, il modello teorico schematizzato e le *propositions* degli autori.

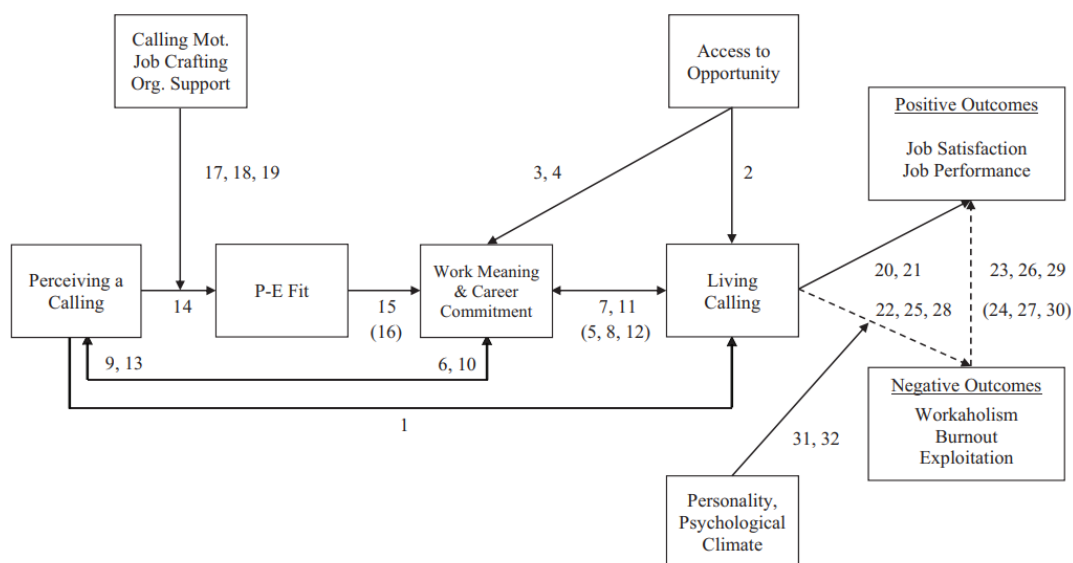


Figura 1.1 Modello teorico di chiamata "Work as Calling Theory" (Duffy, Dik, Velez & Douglass, 2018)

Gli autori individuano come predittori della chiamata professionale la percezione di chiamata (*perceiving a calling*) e l'accesso all'opportunità (*access to opportunity*).

Perceiving a calling. Percepire la chiamata risulta essere un diretto predittore del vivere la chiamata (1) (Duffy et al., 2018). Percepire e vivere la chiamata risultano quindi essere due variabili distinte, con una correlazione misurata dagli studiosi tra .35 e .54 (Duffy et al., 2013; Bott et al. 2012; Duffy et al., 2016). Non è possibile vivere una chiamata se prima essa non è stata percepita (Duffy et al., 2018). È però possibile percepire la chiamata e poi non viverla: in questo caso di *unanswered calling*, gli individui provano rimorso per non aver perseguito il percorso di carriera desiderato (Berg et al., 2010) e risultano provare livelli di insoddisfazione maggiore dal punto di vista lavorativo, di vita e di salute rispetto ad individui che non hanno mai percepito una chiamata (Gazica & Spector, 2015).

Access to opportunity. L'accesso all'opportunità (e la percezione di accesso all'opportunità) risulta essere un predittore diretto del vivere la chiamata (2). La possibilità di inseguire la propria *calling* può essere ostacolata da sistemi di oppressione come classismo, sessismo ed eterosessismo, che limitano le opportunità di occupazione degli individui con status sociale marginalizzato. Questi individui potrebbero accettare lavori non allineati con le proprie aspirazioni e caratteristiche personali per la propria sopravvivenza (Duffy et al., 2016).

La percezione della chiamata e il vivere la chiamata sono mediati da esperienze sul posto di lavoro. In particolare, Duffy et al. affermano come i costrutti di *work meaning* e *career commitment* possano facilitare il collegamento tra la percezione e vissuto di chiamata. Il *work meaning* e il *career commitment* sono predetti dalla percezione di

chiamata (6 e 10) e predicono il vivere la chiamata (7 e 11): questi costrutti mediano quindi la relazione tra *perceiving a calling* e *living a calling* (8 e 12). La relazione tra i due costrutti e la percezione della chiamata risulta essere anche reciproca. Per questo motivo, individui che sviluppano un senso di significato nel lavoro e percepiscono un alto livello di impegno nella propria carriera saranno portati nel tempo a percepire tale carriera come una chiamata (9 e 13). Entrambi i costrutti sono inoltre predetti dall'accesso all'opportunità (3 e 4) e ne mediano la relazione con il vivere la chiamata (5).

Dik e collaboratori (2018) individuano inoltre altre variabili individuali e lavorative che possano connettere il *work meaning* e il *career commitment* alla percezione di chiamata. In particolare, individuano come mediatore di tale relazione il costrutto di *person-environment fit* (P-E fit) e come moderatori tra la percezione di chiamata e il *P-E fit* i costrutti di *calling motivation*, *job crafting* e *organizational support*.

La variabile P-E fit è direttamente predetta dalla percezione di chiamata (14) e predice direttamente *work meaning* e *career commitment* (15). Essa media, e quindi spiega, questa relazione (16): il motivo per cui gli individui che percepiscono una chiamata avvertono più impegno sul lavoro e gli attribuiscono più senso è che essi tendono ad entrare in contatto con ambienti lavorativi adatti a sé.

Le tre variabili mediatrici alterano la relazione tra la percezione di chiamata e il *P-E fit*. La variabile *calling motivation* è relativa all'impegno che l'individuo è disposto a investire nel perseguimento della chiamata, mentre la variabile *job crafting* è relativa alla capacità di modificare il proprio ambiente lavorativo in modo da creare un miglior adattamento. Tutte e tre le variabili, nel loro effetto di mediazione, aumentano la relazione tra la percezione di chiamata e l'adattamento persona-ambiente (17, 18, 19).

Il modello *WCT* evidenzia infine gli esiti derivanti dal vivere la chiamata, distinguendoli in *outcomes* positivi e negativi. Gli esiti positivi riguardano le seguenti ipotesi: vivere la chiamata predice positivamente la soddisfazione sul lavoro e la *job performance* (20 e 21) (Park et al. 2016; Lee et al., 2016). Gli esiti negativi corrispondono a quella che gli autori definiscono come la *dark side of calling*: per alcuni individui, vivere la chiamata può portare ad un maggior livello di *workaholism*, *burnout* e *organizational exploitation* (22, 25, 28). Questi *outcomes* negativi predicono direttamente la diminuzione di *job satisfaction* e *job performance* (23, 26, 29). Essi mediano inoltre la relazione tra il vivere la chiamata e la *job satisfaction* (24, 27, 30). Sono infine presi in considerazione i fattori di personalità individuale e clima psicologico come moderatori tra il vivere la chiamata e gli *outcomes* negativi. In particolare, gli individui che possiedono alcuni tratti maladattivi della personalità come nevroticismo, perfezionismo, bassa autostima e bassa coscienziosità evidenzieranno più alti livelli di *burnout*, *workaholism* e *organizational exploitation* (31 e 32).

1.3 Le diverse misure di chiamata professionale

Le misure del costrutto di chiamata sono state costruite quasi esclusivamente in maniera *ad-hoc*, a seconda dei bisogni individuali di ogni studio (Dik et al., 2012). Spesso, i costrutti operazionalizzati sono definiti solo vagamente e l'evidenza psicometrica approfondita è raramente riportata (Dik et al., 2012). Tipicamente, le strategie di misurazione si sono suddivise in tre categorie: (a) suddivisione dei partecipanti in orientamento per “*calling*” “*job*” or “*career*” basandosi su valutazioni di vignette che descrivono diversi orientamenti lavorativi (ad esempio, Peterson et al., 2009; Wrzesniewski et al., 1997); (b) sviluppo di scale costruite per una specifica popolazione di lavoratori (ad esempio, Bunderson & Thompson, 2009; Dobrow & Tosti-Kharas, 2011;

Serow, 1994); (c) scale multi-item unidimensionali (ad esempio, Dobrow & Tosti-Kharas, 2011; Dreher et. al, 2009). Ognuna di queste modalità di misura del costrutto di chiamata contiene in sé limitazioni (Dik et al., 2012). Le tecniche di categorizzazione sono poco affidabili e possono limitare la capacità di valutare la variabilità del costrutto di chiamata, mentre gli strumenti sviluppati per specifiche occupazioni sono spesso utili nell'immediato contesto della ricerca, ma difficilmente generalizzabili su ulteriori campioni (Dik et al., 2012). Infine, le scale unidimensionali sono limitate nell'esaminare le complessità di come gli individui esprimono e vivono un senso di chiamata: il costrutto è stato infatti concettualizzato come di natura multidimensionale (ad esempio, Dik & Duffy, 2009; Elangovan et al., 2010). Gli studiosi Thompson & Bunderson (2019) hanno effettuato una *review* sulle misure chiave di chiamata, attraverso l'analisi di 84 articoli dal 1997 e il 2018. Negli studi analizzati, i ricercatori riportano la mancanza dell'utilizzo di una definizione standard e, di conseguenza, la mancanza di una misura standard per il costrutto (Thompson & Bunderson, 2019). In Tabella 1.1 si riportano, descritti e comparati, i sette strumenti ricorrenti utilizzati nel campione di ricerche di Thompson e Bunderson (2019).

Strumento	Definizione utilizzata	Descrizione dello strumento
Work Orientation Scale (Wrzesniewski et al. 1997)	Una chiamata è lavoro non separabile dalla vita dell'individuo e motivato con "la soddisfazione che svolgere il lavoro porta all'individuo" (p. 22)	Include una risposta single-item ad un paragrafo che descrive la chiamata e una scala a 7 items
The Brief Calling Scale (Steger & Dik, 2006)	Gli items fanno riferimento alla chiamata ma non è fornita una definizione	Composta da due sotto scale (presenza della calling e ricerca della calling), ognuna composta da 2 items
The Living One's Calling Scale (Duffy et al. 2012a)	Gli items fanno riferimento alla chiamata ma non è fornita una definizione	Scala a 6 items che misura se il rispondente è coinvolto in attività lavorative con un senso di chiamata
Dobrow & Tosti-Kharas' (2011) Calling Scale	La chiamata è "una passione profonda e significativa che	Scala a 12 items

	le persone sperimentano verso un dominio” (p. 1000)	
The Calling and Vocation Questionnaire (Dik et al. 2012)	La chiamata è composta da tre dimensioni: richiamo trascendente, lavoro propositivo e orientamento prosociale	Per ognuna delle tre dimensioni, 4 items misurano la presenza di calling e 4 items misurano la ricerca della chiamata. Totale di 24 items
Neoclassical Calling Scale (Bunderson & Thompson 2009)	La chiamata è “il posto nella visione del lavoro della società che un individuo si sente destinato a ricoprire per un particolar dono o talento e/o opportunità di vita idiosincratiche”	Scala a 6 items
The Calling Motivation Scale (Duffy et al. 2015)	Gli items fanno riferimento alla chiamata ma non è fornita una definizione	Strumento a 3 items che misura l'intensità con cui un individuo è spinto a perseguire una chiamata

Tabella 1.1 Misure del costrutto di calling, da Thompson & Bunderson, 2019

1.3.1 Unified Multidimensional Calling Scale (UMCS)

Vianello et al. (2018) propongono uno strumento per la misura della chiamata, la *Unified Multidimensional Calling Scale*, con lo scopo di combinare la prospettiva teorica neoclassica e quella moderna. Il modello testato dimostra che il costrutto di *calling* è multidimensionale, e che le diverse prospettive di chiamata fanno riferimento allo stesso costrutto (Vianello et al., 2018).

Lo strumento teorizzato e testato si compone di 22 items, che fanno riferimento a sette *facets*: passione, significato, sacrificio, pervasività, orientamento prosociale, evocazione trascendente e identità. Queste sette dimensioni identificate sintetizzano le due prospettive del costrutto di chiamata, e sono presenti nei principali quadri teorici affrontati in precedenza. In Tabella 1.2 (Vianello et al., 2018) vengono riportati i *facets* e le rispettive teorie di *calling* in cui essi sono presenti.

	Dobrow e Tosti-Kharas (2011)	Dik Duffy et al. (2009, 2012)	Wrzesniewski et al. (1997)	Bunderson & Thompson (2009)	Paskova et al. (2015)	Hagmeier & Abele (2012)
Identità	X		X			X
Pervasività	X		X			
Significato	X	X			X	
Evocazione trascendente		X		X	X	X
Orientamento prosociale		X	X		X	X
Sacrificio	X					
Passione	X		X		X	

Tabella 1.2 Le sette dimensioni di calling in sei diverse prospettive teoriche, da Vianello et al., 2018

Gli items del modello *UMCS* sono estratti dalle due scale più popolari di calling: la *Calling Scale* di Dobrow & Tosti-Kharas (2011) e il *Calling and Vocation Questionnaire* (CVQ) di Dik e colleghi (2012). La scala di Dobrow & Tosti-Kharas (2011) presenta 12 items che fanno riferimento ai domini di passione, sacrificio e identità. Vianello et al. aggiungono un tredicesimo item relativo al dominio di sacrificio per bilanciare il numero di items nei *facets*. (Vianello et al., 2018). Il CVQ di Dik e colleghi (2012) è composto da 24 items che fanno riferimento alle tre sotto scale di evocazione trascendente, lavoro significativo e orientamento prosociale. Da questa scala, Vianello e colleghi hanno selezionato i nove items con i più alti loading sui rispettivi fattori.

Il modello *UMCS* si è dimostrato valido e affidabile in due studi condotti su studenti universitari (N= 5886) e lavoratori (N= 205), confermando la presenza dei sette *facets* nel

costrutto di chiamata. La scala risulta inoltre invariante nel tempo e tra i vari domini di *calling*. La scala UMCS-22 è stata successivamente integrata con sei items, nella scala UMCS-28 (Vianello et al., 2022; Gerdel et al., 2022). La scala UMCS-28 risulta avere una più alta coerenza interna della precedente e una migliore struttura fattoriale (Vianello et al., 2022).

2. L'INVARIANZA DELLA MISURA

2.1 I problemi di comparabilità nella ricerca cross-culturale

Van de Vijver (1998) definisce ricerca cross-culturale qualsiasi studio comparativo che coinvolga diversi stati nazionali o diversi gruppi culturali all'interno di un singolo paese. Negli ultimi decenni la ricerca cross-culturale ha subito una forte crescita, raccogliendo sempre maggiore interesse tra i ricercatori (Kankaras & Moors, 2010). Studi cross-nazionali, cross-culturali e multilinguistici sono utilizzati per comparare opinioni, attitudini, valori e abilità tra diversi gruppi culturali. Questo crescente interesse per gli studi comparativi può essere spiegato dai rapidi cambiamenti che stanno avvenendo nella nostra società (Van de Vijver, 1998): la globalizzazione economica, sociale, politica e tecnologica e i sempre crescenti flussi migratori stanno abbattendo confini nazionali secolari, trasformando fenomeni locali e regionali in fenomeni globali. In Europa, i processi di integrazione e l'espansione dell'Unione Europea hanno ulteriormente aumentato la produzione di ricerca comparativa (Kankaras & Moors, 2010). Nonostante il considerevole interesse per gli studi cross-culturali, continua a non esistere un modo generalmente accettato per trattare alcune problematiche specifiche di questa modalità di ricerca (Van de Vijver, 1998; Kankaras & Moors, 2010). Ad esempio, si è spesso osservato che molti ricercatori assumono che i risultati ottenuti dalla somministrazione dello stesso strumento (ad es. questionari, inventari, test) ai diversi gruppi coinvolti siano automaticamente comparabili (Kankaras & Moors, 2010). Tuttavia, poiché ogni contesto culturale riflette una costellazione di molti fattori, processi e attributi, lo stesso insieme di domande o compiti può assumere significati diversi per individui appartenenti a gruppi culturali diversi, e di conseguenza misurare costrutti diversi in ogni cultura (Kankaras &

Moors, 2010). In questi casi, confrontare semplicemente le medie di scala attraverso i gruppi potrebbe pregiudicare la validità delle conclusioni. Risulta fondamentale in ogni ricerca cross-culturale preoccuparsi di garantire l'equivalenza (i.e. la comparabilità) durante la verifica delle differenze cross-culturali (Hui & Triandis, 1985). Questa comparabilità è chiamata equivalenza o invarianza della misura.

Secondo Van De Vijver (1998) risulta innanzitutto fondamentale stabilire un quadro teorico che tenti di incorporare gli aspetti specifici della ricerca cross-culturale. L'autore indica i costrutti di *bias* e di equivalenza come nucleo di questo quadro teorico, in grado di guidare i piani d'azione in tutte le fasi della ricerca, nello stesso modo in cui i costrutti di validità e affidabilità sono alla base di molte decisioni prese nella ricerca intraculturale (Van De Vijver, 1998).

2.1.1 Il costrutto di *bias* nella ricerca cross-culturale

Il *bias* si presenta nella ricerca cross-culturale quando i risultati ottenuti rappresentano sistematicamente in modo errato i punteggi veri del costrutto analizzato. Il termine *bias* può essere utilizzato genericamente per indicare tutte le questioni legate alla validità specifiche della ricerca cross-culturale. Uno strumento è quindi *biased* (o distorto) se i suoi punteggi non hanno lo stesso significato psicologico tra i gruppi culturali interessati dalla ricerca, e più precisamente se affermazioni sui suoi punteggi (sulle somiglianze e differenze degli stessi) non si applicano nel dominio psicologico dei punteggi (Van De Vijver, 1998). Ad esempio, differenze individuali in test di intelligenza possono riflettere differenze di intelligenza in un singolo gruppo culturale, mentre differenze intergruppo possono essere dovute alla differenza nell'educazione e all'esperienza con i test stessi. Secondo Van de Vijver (1998) da questa definizione del

costrutto di *bias* derivano tre caratteristiche. In primo luogo, i *bias* si riferiscono a fonti involontarie di variazione, e costituiscono una spiegazione alternativa alle differenze intergruppi. Se sono presenti *bias*, le differenze di punteggio in ricerche cross-culturali non sono generate dal costrutto indagato, (ad es. il costrutto di intelligenza o di affiliazione politica), ma da altre caratteristiche (ad es. desiderabilità sociale o educazione). In secondo luogo, i *bias* non sono intrinseci allo strumento utilizzato, ma sono caratteristiche dovute alla specifica comparazione cross-culturale: per questo motivo, la presenza di *bias* potrà essere influenzata sia da caratteristiche dello strumento che da caratteristiche del campione. Infine, i *bias* sono una fonte di variazione sistematica, che è replicabile tra strumenti paralleli somministrati agli stessi campioni (Van de Vijver, 1998). Van de Vijver e Leung (1997) distinguono tre tipologie di bias, elencate in Tabella 2.1.

		Tipo di <i>bias</i>	Descrizione
Bias di metodo		<i>Bias</i> di costrutto	Differenza di costrutto
		<i>Bias</i> di campione	Incomparabilità di campioni
		<i>Bias</i> di strumento	Caratteristiche dello stimolo che inducono differenze cross-culturali (ad es. familiarità dello stimolo)
		<i>Bias</i> di amministrazione	Aspetti procedurali (ad es. problemi di comunicazione)

Tabella 2.1 Tipologie di bias, da Van de Vijver & Leung, 1997

La prima tipologia di *bias* è il *bias* di costrutto, che è presente se il costrutto misurato nello studio non è lo stesso tra le diverse culture. Esso si verifica se un costrutto è definito diversamente o se si sovrappone solo parzialmente tra i gruppi culturali indagati. Ad esempio, Uchida et al. (2004) analizzano il costrutto di felicità tra culture Occidentali e Orientali, evidenziando come in occidente la felicità tenda ad essere definita come un risultato personale, mentre per le culture orientali essa sia concettualizzata come una connessione interpersonale (Uchida et al., 2004). Il *bias* di costrutto aumenta con l'ampliarsi della distanza tra le culture (Kankaras & Moors, 2010). Nel caso in cui un ricercatore sia interessato a comparare attraverso diverse culture un costrutto suscettibile a *bias*, è necessario vengano definiti chiaramente i comportamenti inclusi nella misura presa in esame (Van de Vijver, 1998).

La seconda tipologia di *bias* è il *bias* di metodo, e include tutte le fonti di bias che derivano dagli aspetti metodologici e procedurali dello studio. Questa tipologia di *bias* può essere suddivisa ulteriormente in tre sottotipi. Il primo è il *bias* di campione, e ingloba tutte le differenze di punteggio che sono dovute ad aspetti specifici del campionamento. Si definisce come l'incomparabilità dei campioni dovuta a variazioni di caratteristiche cross-culturali su aspetti esterni rispetto alla variabile ricercata: ne sono un esempio le differenze di background formativo tra culture, che possono influenzare e confondere differenze reali tra i campioni nella somministrazione di test cognitivi (Van de Vijver &

Leung, 1997). La seconda tipologia è il *bias* di strumento, che riguarda errori sistematici che derivano da caratteristiche dello strumento e come gruppi culturali diversi reagiscano ad esso in modo diverso. Questo tipo di *bias* include differenze nella familiarità con lo stimolo, specialmente importante nei test cognitivi, e differenze nella desiderabilità sociale e negli stili di risposta, importante per i test di personalità (Kankaras & Moors, 2010). La terza e ultima tipologia di *bias* di costruito è il *bias* di somministrazione, che è scatenato da problemi comunicativi (ad es. poca comprensione della lingua del test somministrato), caratteristiche dell'intervistatore (ad es. ad sesso e gruppo culturale di appartenenza) e altri aspetti procedurali della raccolta di dati.

L'ultima tipologia di *bias* è il *bias* di item (o Differential Item Functioning; Van de Vijver, 1998). Esso si riferisce ad anomalie dello strumento a livello degli item, quando un item ha differenti significati psicologici attraverso le culture (Van de Vijver & Leung, 1997). Un item di una scala si definisce *biased* se individui appartenenti a culture diverse, che presentano lo stesso punteggio per un tratto latente, non hanno la stessa probabilità di emettere la stessa risposta allo stesso item (Shepard, Camilli & Averill, 1981; Van de Vijver & Leung, 1997). In altre parole, per individui appartenenti a gruppi culturali diversi con gli stessi punteggi totali (i.e. individui di gruppi culturali diversi che sono egualmente intelligenti)) un item *unbiased* dovrebbe essere ugualmente difficile (Van de Vijver & Leung, 1997). L'item *bias* può verificarsi a causa di una traduzione errata, per inapplicabilità dei contenuti di un item in culture differenti o da item che attivano altri tratti o che contengono parole con connotazione ambigua (Van de Vijver & Leung, 1997).

2.1.2 Il costrutto di equivalenza nella ricerca cross-culturale

Van de Vijver (1998) descrive l'equivalenza come il livello di misura a cui una comparazione cross-culturale può essere fatta. Se la misura risulta *biased* verso alcuni gruppi culturali, le differenze individuali all'interno di una cultura e attraverso culture diverse non sono misurate con la stessa scala. (Van de Vijver, 1998). Così come i *bias*, anche l'equivalenza non è intrinseca ad uno strumento, ma caratteristica di una specifica comparazione cross-culturale. Van de Vijver descrive tre livelli di equivalenza elencati in Tabella 2.2.

Tipo di equivalenza	Descrizione
Inequivalenza di costrutto	Differenza di costrutti
Equivalenza di costrutto	Stesso costrutto misurato in ogni gruppo culturale
Equivalenza di unità di misura	Stessa scala (unità di misura) con origini diverse per ogni gruppo culturale
Equivalenza scalare	Stessa scala con la stessa origine in ogni gruppo culturale

Tabella 2.2 Tipi di equivalenza e la loro descrizione, da Van de Vijver (1998)

Le tipologie di equivalenza elencate sono presentate in un ordine gerarchico. La prima, equivalenza di costrutto, si riferisce alla comparabilità di costrutti tra culture (i.e., quando si comparano erroneamente “mele e arance” esiste non-equivalenza di costrutto,

Van de Vijver & Tanzer, 1997). Essa è la tipologia più debole di equivalenza, chiamata anche equivalenza funzionale o equivalenza strutturale (Van de Vijver, 1998), e si verifica quando lo stesso costrutto è stato misurato tra gruppi culturali, non necessariamente con l'utilizzo dello stesso strumento. La seconda tipologia elencata da Van de Vijver è l'equivalenza di unità di misura: essa si verifica se l'unità di misura è identica tra i gruppi, mentre l'origine è diversa. Questo tipo di equivalenza si verifica, ad esempio, se la temperatura viene misurata utilizzando sia la scala Celsius che la scala Kelvin: le unità di misura sono identiche ma presentano una differenza costante (i.e. uno scostamento) di 273 gradi tra le misure. Le differenze individuali possono essere misurate a livello di rapporto in ogni gruppo mentre non è possibile alcuna comparazione tra culture. Infatti, a differenza dell'esempio sulla temperatura, non è quasi mai possibile conoscere lo scostamento tra le misure nelle scienze sociali e comportamentali (Van de Vijver, 1998). L'ultima equivalenza descritta da Van de Vijver è l'equivalenza scalare, o comparabilità *full score*. In questo caso, lo stesso livello intervallo o rapporto si applica alle misure tra le culture comparate. Questo è il tipo di equivalenza che viene ipotizzato quando le medie sono comparate tra culture, come ad esempio nei *t* test e analisi della varianza e in generale in tutti i test in cui si assume che le varianze tra i due gruppi siano uguali.

Secondo Van de Vijver, tipologie più alte di equivalenza sono meno robuste contro i *bias*: l'equivalenza scalare sarà quindi più suscettibile a bias dell'equivalenza di unità di misura. Il ricercatore afferma infine che solo in presenza del livello di equivalenza scalare e assumendo l'assenza di ogni *bias* è possibile fare affermazioni come “la Cultura A ha un punteggio più alto su una determinata misura della Cultura B” (Van de Vijver, 1998).

2.2 Il costrutto di misura secondo la Teoria Classica dei Test

La misura può essere definita come l'assegnazione sistematica di numeri su variabili per rappresentare caratteristiche di persone, oggetti o eventi. Nelle scienze organizzative, i processi di misura sono tipicamente indirizzati a descrivere caratteristiche di individui, gruppi o organizzazioni che risultano essere di interesse per i ricercatori (Vandenberg & Lance, 2000). La misura continua ad essere di cruciale interesse per la ricerca psicologica: ne sono evidenza le numerose riviste scientifiche dedicate alla questione della misura (ad es., *Structural Equation Modeling*, *Multivariate Behavioral Research*) e l'ampio spazio dedicato alle questioni riguardanti la misura nelle sezioni di metodo e discussione negli articoli (Vandenberg & Lance, 2000). Storicamente, la valutazione della qualità della misura è stata legata alla teoria classica dei test (CTT; Crocker & Algina, 1986; Lord & Novick, 1968; Nunnally & Bernstein, 1994).

Uno degli aspetti fondamentali della CTT è la definizione dei punteggi osservati in termini di componenti di punteggio vero e di errore:

$$X_{ijk} = T_{ij} + E_{ijk} \quad \text{Eq. 1}$$

Dove X_{ijk} si riferisce alla k osservazione del punteggio di i sulla misura del tratto j , T_{ij} è il punteggio vero di i sul tratto j , e E_{ijk} rappresenta una componente di errore di misura non sistematico che, teoricamente, è campionato da una distribuzione normale con media zero e σ_E^2 . T_{ij} è definito come il valore atteso tra le osservazioni di X_{ijk} , e si assume che T_{ij} non sia correlato con E_{ijk} . Quest'ultima assunzione permette la scomposizione della varianza del punteggio osservato (σ_X^2) nelle componenti di punteggio vero ed errore, o

$$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$$

Eq. 2

L'equazione afferma che la varianza dei punteggi individuali osservati riflette la varianza contribuita dalle differenze individuali vere nel tratto misurato e la varianza contribuita dall'errore di misura non sistematico. La definizione del test di attendibilità in termini di $r_{XX'} = \sigma_T^2 \div (\sigma_T^2 + \sigma_E^2) = \sigma_T^2 \div \sigma_X^2$ deriva direttamente dall'Equazione 2 ed è un criterio per la valutazione delle proprietà psicometriche dei test, che deriva dall'aspettativa che X_{ijk} rifletta massimamente il punteggio vero (T_{ij}) escludendo l'errore (E_{ijk}). La Figura 2.1 mostra come le componenti della CTT e le connessioni tra costrutti latenti e operazioni impiegate per rappresentarli sono strettamente connesse (Vandenberg & Lance, 2000).

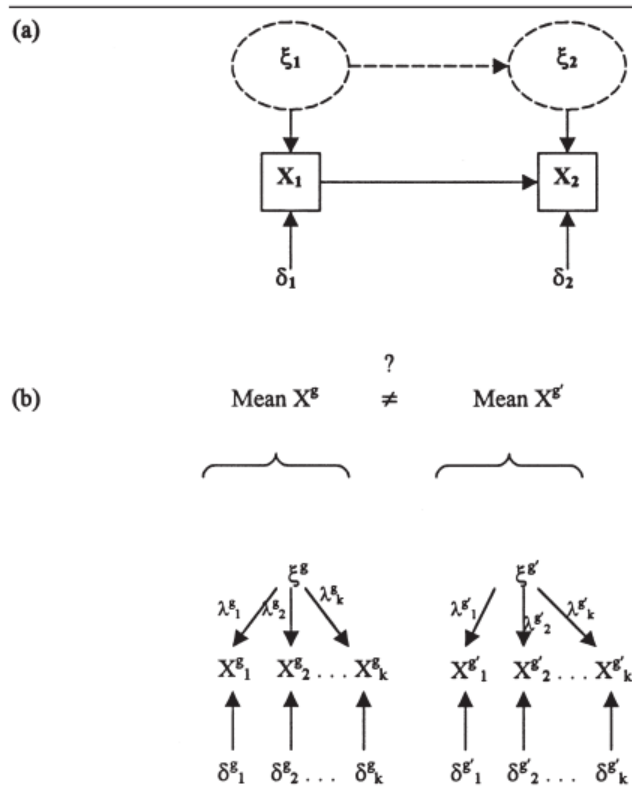


Figura 2.1 Collegamento tra punteggi osservati e veri (a) e il modello di misura implicato sottostante a comparazioni semplici tra gruppi (b), da Vandenberg & Lance, 2000

Nella Figura 2.1 (a), ξ è il costrutto ipotetico di interesse, corrispondente a T_{ij} . Il cerchio tratteggiato simbolizza come ξ sia una variabile latente, così come il punteggio vero definito in termini di valore atteso di realizzazione su X_{ijk} . La freccia che va in direzione da ξ a X riflette l'idea che la variazione nella variabile latente causi una variazione in X , e che i sistemi di misura siano imperfetti: una parte della variazione di X può essere attribuita ad influenze esterne, inclusi errori sistematici e non sistematici (E_{ijk} in Equazione 1). La Figura 2.1 (b) presenta una semplice comparazione a due gruppi con lo scopo di spiegare fino a che punto le proprietà della misura delle variabili manifeste (i.e. X) siano generalizzabili tra popolazioni (Vandenberg & Lance, 2000). È necessario assumere l'esistenza di due gruppi a cui è somministrata la misura di interesse. Sono successivamente formati punteggi composti e le medie dei gruppi sui punteggi composti sono comparate statisticamente. Per poter fare inferenze e trarre conclusioni sulle differenze tra gruppi, gli autori che applicano la CTT fanno tipicamente affermazioni sulla attendibilità e validità delle misure, il cui principale scopo è quello di supportare la credenza che la misura manifesta X rifletta, come atteso, l'attributo latente sottostante ξ . Ciononostante, come indicato nella parte inferiore della Figura 2.1 (b) (al di sotto di “{“), il modello di misura sottostante alle medie osservate è più complesso di quanto si possa attendere. ξ^g implica un dominio di contenuto nel gruppo g che può essere solo rappresentato adeguatamente utilizzando operazionalizzazioni multiple (Embretson, 1983), e le risposte nella k operazionalizzazione nel gruppo g (X_k^g) riflettono in parte il costrutto teorico voluto, ma in parte altri fattori unici non spiegati che includono ma non si limitano all'errore di misura non sistematico (Vandenberg & Lance, 2000). Secondo gli autori, evidenze di attendibilità e validità rimangono importanti per valutare gli attributi

psicometrici di una misura, ma una interpretazione non ambigua delle medie osservate dipende anche dall'equivalenza tra gruppi del modello di misura sottostante (Vandenberg & Lance, 2000). Nello specifico, una comparazione tra gruppi (i.e., t test, ANOVA) con buona attendibilità e validità delle misure, richiede comunque implicitamente presupposti non testati: equivalenza concettuale della variabile teorica sottostante ξ in ogni gruppo, associazioni equivalenti (λ_k^g) tra operationalizzazioni (X_k^g) e ξ tra i gruppi e, infine, la misura in cui le operationalizzazioni (X_k^g) sono influenzate allo stesso modo e forse dagli stessi fattori unici δ_k^g tra i gruppi. Questi presupposti sono raramente testati, ma la loro violazione può rendere molto sospette le interpretazioni dei confronti tra gruppi sulle misure non equivalenti (Vandenberg & Lance, 2000). Ad esempio, la prima *assunzione* (punteggi veri o costrutti teorici equivalenti) presuppone che le misure manifeste evocino lo stesso quadro di riferimento in ognuno dei gruppi comparati: nel caso in cui le misure significhino cose differenti per gruppi differenti, una comparazione tra medie dei gruppi potrebbe essere fuorviante e condurre a conclusioni errate. Come affermato da Vandenberg e Lance (2000), comparazioni attraverso i gruppi richiedono presupposti di invarianza della misura tra i gruppi in esame. Pertanto, la dimostrazione dell'invarianza delle misure è un prerequisito logico per la valutazione di ipotesi sostanziali sulle differenze di gruppo, indipendentemente dal fatto che il confronto sia semplice come un test sulle differenze medie tra i gruppi o complesso come la verifica dell'invarianza di un modello strutturale teorico tra i gruppi. (Vandenberg & Lance, 2000).

2.3 Il costrutto di invarianza della misura

Secondo Horn e McArdle (1992, p. 117) la domanda generale a cui risponde l'invarianza di misura è il capire “se, in condizioni diverse di osservazione e studio dei fenomeni, le operazioni di misurazione producano o meno misure dello stesso attributo”. Secondo i ricercatori, nel caso in cui non ci sia evidenza che indichi la presenza o l'assenza di invarianza della misura, o nel caso in cui ci sia evidenza dell'assenza di invarianza, le basi per trarre inferenze scientifiche risultano carenti: differenze trovate tra individui o tra gruppi non possono essere interpretate senza ambiguità (Horn & McArdle, 1992). Vandenberg & Lance (2000) affermano che il costrutto di invarianza della misura trova un sinonimo nel termine equivalenza della misura.

Nel loro articolo “*A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research*” i ricercatori Vandenberg e Lance (2000) analizzano problematiche legate alla misura che non risultano incluse nella portata della Teoria Classica dei Test (CTT), in particolare il costrutto di invarianza della misura. Secondo i ricercatori, le questioni sottostanti alla invarianza o equivalenza della misura che non sono direttamente indirizzabili con modalità utilizzate dalla CTT sono le seguenti: (a) gli individui appartenenti a culture differenti interpretano una misura in una maniera concettualmente simile? (b) Le fonti di classificazione definiscono la performance in modo simile quando valutano lo stesso soggetto su dimensioni di performance identiche? (c) ci sono differenze di genere, etniche o individuali che precludono l'individuo dal rispondere agli strumenti in modo simile? (d) il processo stesso di interesse sostanziale (i.e., un intervento o una manipolazione sperimentale) altera il quadro concettuale di riferimento rispetto al quale un gruppo risponde a una misura nel tempo?

I ricercatori sintetizzano nella Figura 2.1 (b), esaminata nel paragrafo 2.2, la questione riguardante l'invarianza della misura. La Figura è in realtà una semplificazione della questione dell'equivalenza di misura e rappresenta solo un sottoinsieme degli attributi di misura per i quali è necessario ipotizzare l'equivalenza di misura tra gruppi quando si effettuano confronti tra gruppi. Le relazioni tra misure tra i k items nel gruppo g in Figura 2.1 (b) possono essere rappresentati come

$$X_k^g = \tau_k^g + \Lambda_k^g \xi^g + \delta_k^g \quad \text{Eq. 3}$$

Dove X_k^g si riferisce al vettore di item comprendente la misura composita, Λ_k^g si riferisce alla matrice del coefficiente angolare della regressione che mette in relazione X_k^g e ξ^g , τ_k^g si riferisce al vettore dell'intercetta di regressione e δ_k^g si riferisce al vettore di fattori unici. Assumendo che $E(\xi^g, \delta_k^g) = 0$, l'equazione di covarianza che segue dall'Equazione 1 è

$$\Sigma^g = \Lambda_X^g \Phi^g \Lambda_X^{g'} + \Theta_\delta^g \quad \text{Eq. 4}$$

Dove Σ^g rappresenta la matrice delle varianze e covarianze tra i K items nella popolazione g (il gruppo), Λ_X^g è la matrice della saturazione degli item su ξ^g , Φ^g contiene varianze e covarianze tra ξ^g , e Θ_δ^g è la matrice diagonale di varianze fra le unicità. Le Equazioni 3 e 4 dimostrano come l'invarianza della misura sia testabile con un quadro teorico di Analisi Fattoriale Confermativa (CFA) che sarà trattata nel corso del capitolo.

2.3.1 L'invarianza della misura configurale, metrica, scalare e stretta

Da un punto di vista della misura, esistono livelli di invarianza o equivalenza gerarchicamente connessi che assumono livelli sempre più forti di comparabilità di misura tra culture. I livelli più importanti di equivalenza della misura sono configurale, metrico, scalare e stretto (Vandenberg & Lance, 2000). I livelli sono ordinati gerarchicamente, e i livelli di equivalenza più alti presuppongono quelli inferiori (Kankaras & Moors, 2010). Livelli di equivalenza più alti sono più difficili da ottenere dato che forniscono un test più forte di invarianza cross-culturale, ma permettono una forma più estesa di comparazioni cross-culturali o cross-temporali.

Per una rappresentazione grafica dei livelli di invarianza, sarà utilizzato un esempio illustrativo tratto dall'articolo "*Testing measurement invariance across groups: Applications in cross-cultural research*" di Milfont & Fischer (2010). I due ricercatori utilizzano la rappresentazione grafica tipica dei modelli di Analisi Fattoriale Confermativa: cerchi o ovali per rappresentare le variabili latenti non osservabili e quadrati per rappresentare le variabili manifeste e osservate. Le linee sono utilizzate per indicare le relazioni tra variabili: linee con una sola freccia che si rivolgono verso una specifica variabile rappresentano una relazione diretta ipotizzata, dove la variabile verso cui punta la freccia rappresenta la variabile dipendente. Linee con frecce da entrambe le estremità indicano che le variabili sono collegate tra loro, senza una direzione presunta della relazione. In Figura 2.2 è mostrato il modello esemplificativo utilizzato: sono presenti quattro variabili osservate (i.e., READING 5, WRITING 5, READING 7, WRITING 7) e due variabili latenti (i.e., Verbal 5, Verbal 7).

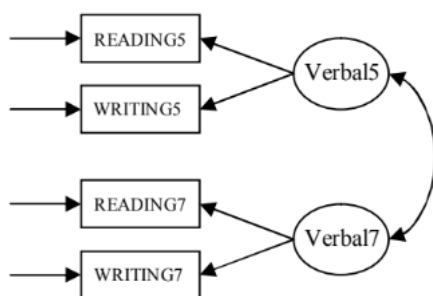


Figura 2.2, modello teorico utilizzato nell'esempio illustrativo, da Milfont & Fischer, 2010.

Il primo livello di invarianza, invarianza *configurale*, è definito come *lo stesso numero di fattori e lo stesso pattern di saturazioni libere o fissate tra gruppi*, senza restrizioni di uguaglianza in nessun altro parametro del modello. L'invarianza configurale è una condizione necessaria per testare l'invarianza di parametri di misura (ad es., saturazioni, intercette degli item, varianza d'errore) che riguardano le caratteristiche di misura degli item o indicatori osservati (Wang & Wang, 2012). Se l'invarianza configurale non è dimostrata, gli indicatori osservati della scala sotto esame misurano costrutti differenti in gruppi differenti, perciò test successivi sono immotivati (Wang & Wang, 2012). Questo livello di invarianza è anche indicato in letteratura come “*baseline model*” (Bagozzi & Edwards, 1998; Marsh, 1994; Reise et al., 1993), test di uguaglianza di strutture fattoriali (Cole & Maxwell, 1895), o uguaglianza numerica di fattori e pattern di fattori (Taris et al., 1998). L'invarianza configurale è illustrata in Figura 2.3 (Milfont & Fischer, 2010).

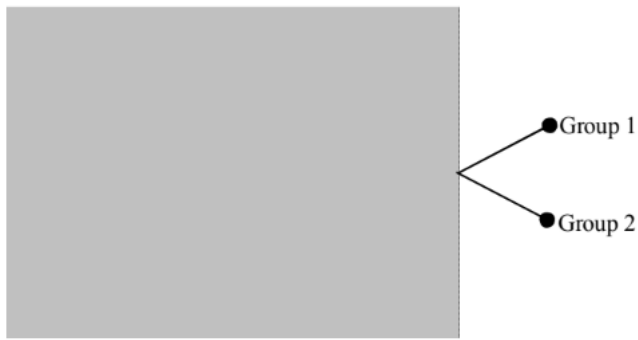


Figura 2.3, Invarianza configurale (stessa struttura tra i gruppi), da Milfont & Fischer, 2010.

Nota: le sezioni di colore grigio rappresentano parti del modello in Figura 2.2 non interessate dall'invarianza configurale

Il secondo livello di invarianza è l'invarianza *metrica*. Essa è definita come *invarianza delle saturazioni* tra i gruppi. Le saturazioni sono i coefficienti angolari della regressione degli item o indicatori osservati sul fattore sottostante, e rappresentano la forza della relazione lineare (Wang & Wang, 2012). Se le saturazioni sono invarianti tra i gruppi (i.e., $\Lambda^g = \Lambda^{g'}$, Vandenberg & Lance, 2000), allora le misure tra i gruppi sono considerate essere sulla stessa scala, di conseguenza l'invarianza delle saturazioni è anche detta invarianza metrica (Wang & Wang, 2012). Quando l'invarianza delle saturazioni è dimostrata, le variabili o fattori latenti sottostanti sono misurati nello stesso modo in tutti i gruppi; perciò, continuare a testare l'invarianza della varianza fattoriale (i.e., la dispersione dei punteggi di fattori) e l'invarianza della covarianza fattoriale (i.e., le relazioni tra fattori) è significativo (Wang & Wang, 2012). Questo livello di invarianza è inoltre indicato in letteratura come test di pattern fattoriali invarianti (Alwin & Jackson, 1981), uguaglianza delle unità di scala (Cole & Maxwell, 1985; Schaie & Hertzog, 1985), comparabilità metrica (Drasgow & Kanfer, 1985), invarianza fattoriale (Horn & McArdle, 1992), invarianza dei loading dei fattori (Marsh, 1994), *full measurement*

invariance (Reise et al., 1993) o *weak measurement invariance* (Wang & Wang, 2012).

L'invarianza metrica è illustrata in Figura 2.4 (Milfont & Fischer, 2010).

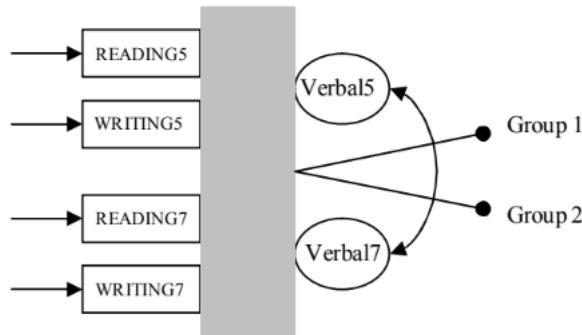


Figura 2.4, Invarianza metrica (stesse saturazioni tra i gruppi), da Milfont & Fischer, 2010

Nota: le sezioni di colore grigio rappresentano parti del modello in Figura 2.2 non interessate dall'invarianza metrica

Il terzo livello di equivalenza è l'invarianza *scalare* della misura. Esso è definito come invarianza sia delle saturazioni che delle *intercette* degli item o indicatori tra i gruppi (Wang & Wang, 2012). L'intercetta dell'item rappresenta l'origine (o scala) della misura. Se le intercette degli item non sono invarianti tra i gruppi, i partecipanti di almeno uno dei gruppi tendono a rispondere in modo sistematicamente più elevato o inferiore agli item della scala sotto studio. (Wang & Wang, 2012). Quando l'invarianza è confermata sia per le saturazioni che per le intercette degli item (i.e., $\tau^g = \tau^{g'}$, Vandenberg & Lance, 2000), i punteggi agli item di gruppi differenti hanno la stessa metrica di misurazione e la stessa grandezza scalare: i fattori sottostanti possono quindi essere comparati tra i gruppi (Wang & Wang, 2012). Questo livello di invarianza è inoltre indicato in letteratura come *strong measurement invariance* (Wang & Wang, 2012). L'invarianza scalare è illustrata in Figura 2.5 (Milfont & Fischer, 2010).

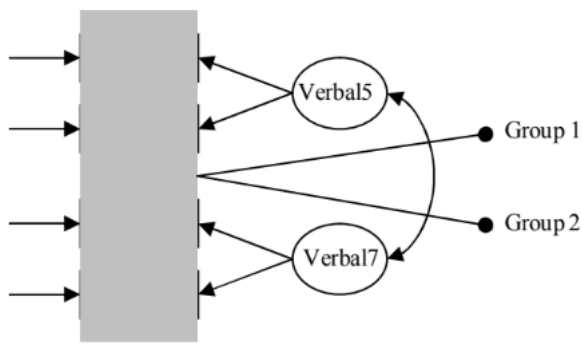


Figura 2.5, invarianza scalare (stesse intercette degli item tra i gruppi), da Milfont & Fischer, 2010

Nota: le sezioni di colore grigio rappresentano parti del modello in Figura 2.2 non interessate dall'invarianza scalare

Il quarto e ultimo livello di equivalenza è l'invarianza *stretta* della misura. Essa richiede invarianza metrica (i.e., saturazioni), scalare (i.e., intercette dell'indicatore o item) e della *varianza d'errore* (Wang & Wang, 2012). L'invarianza della varianza d'errore è considerata come l'invarianza dell'attendibilità degli item tra i gruppi. Essa è raramente utile nella pratica, e testarla è considerato non necessario (Bentler, 2005). Sebbene questo livello di invarianza possa non essere testato quando si analizzano le variabili latenti, gli item utilizzati in una misura composita devono soddisfare lo standard di invarianza stretta per evitare conclusioni errate dovute a cambiamenti nell'affidabilità (Newsom, 2015). L'invarianza stretta è illustrata in Figura 2.6 (Milfont & Fischer, 2010).

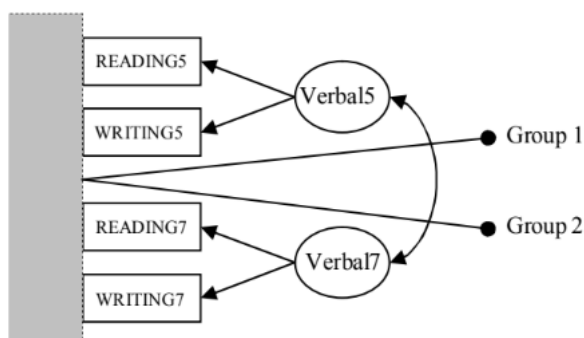


Figura 2.6, invarianza stretta (stessa varianza d'errore tra i gruppi), da Milfont & Fischer, 2010.

Nota: le sezioni di colore grigio rappresentano parti del modello in Figura 2.2 non interessate dall'invarianza stretta

Il modello di invarianza della misura appena descritto si riferisce all'invarianza della misura piena o completa, in quanto stabilisce se ogni elemento dato (i.e., le saturazioni, l'intercetta degli item, la varianza dei fattori) risulta uguale in tutti i gruppi. Tuttavia, è improbabile che la piena invarianza della misura si verifichi nella pratica. Per ovviare all'obiettivo troppo rigido e irrealistico che i vincoli di invarianza debbano valere per tutti i parametri tra i gruppi, Byrne et al. (1989) introducono il concetto di invarianza parziale della misura, in cui solo un sottoinsieme di parametri del modello è vincolato a essere invariante, mentre un altro sottoinsieme di parametri può variare tra i gruppi. Pertanto, l'invarianza parziale della misura può consentire confronti appropriati tra i gruppi anche se non si ottiene l'invarianza completa della misura. L'invarianza parziale della misura può essere valutata in due casi: nel caso in cui le misure siano invarianti tra alcuni ma non tra tutti i gruppi, e nel caso in cui alcuni ma non tutti i parametri siano invarianti tra i gruppi (Vandenberg & Lance, 2000).

Come precedentemente affermato l'invarianza della misura è testabile con un quadro teorico di Analisi Fattoriale Confermativa (CFA). In particolare, l'approccio metodologico prevalente per la valutazione dell'invarianza della misura è l'Analisi Fattoriale Confermativa Multi-Gruppo (MGCFA) (Byrne et al., 1989; Meredith, 1993; Vandenberg & Lance, 2000).

2.4 L'Analisi Fattoriale Confermativa (CFA)

In statistica e in psicometria l'analisi fattoriale è una tecnica che permette di evidenziare l'esistenza di una struttura di tratti latenti (in psicometria) o fattori o dimensioni (in statistica), non misurabili direttamente, all'interno di un insieme di

variabili direttamente osservabili che si relazionino con tali tratti latenti (Gallucci et al., 2017). L'obiettivo di una analisi fattoriale è quello di spiegare la variabilità esistente tra una serie di variabili manifeste direttamente osservate ed i loro legami in termini di un numero ridotto di variabili latenti: i fattori. In termini più formali: date p variabili manifeste osservate su n individui $x_1 \dots x_p$, l'obiettivo di una analisi fattoriale è di riformulare le p variabili manifeste in termini di q fattori “comuni” (con $q < p$) e p errori:

$$x_p = \lambda_{p1}\xi_1 + \lambda_{p2}\xi_2 + \dots + \lambda_{pq}\xi_q + \delta_p \quad \text{Eq. 5}$$

Dove λ_p indica la saturazione, ξ_1 i fattori comuni e δ_p il fattore unico.

L'analisi fattoriale può essere condotta sia a scopi esplorativi (EFA, Exploratory Factor Analysis) che a scopi confermativi (CFA, Confirmatory Factor Analysis): nel primo caso i fattori vengono estrapolati a partire dai dati, mentre nel secondo è il ricercatore a porre dei vincoli sul proprio modello e a verificare se tale modello sia coerente con i dati osservati, cioè, in termini statistici, se esso mostra un *fit* adeguato ai dati (Furr & Bacharach, 2013). L'analisi fattoriale confermativa si concentra sul modellamento della relazione tra gli indicatori manifesti (i.e., osservati) e le variabili latenti sottostanti, i fattori.

Essa viene condotta attraverso l'uso di modelli di equazioni strutturali (SEM, Structural Equation Modeling), un approccio di analisi che combina la *path analysis* e l'analisi fattoriale confermativa (Weston & Gore, 2006). L'obiettivo del SEM è simile a quello dell'analisi fattoriale: fornire una sintesi parsimoniosa delle interrelazioni tra le variabili. Esso è anche affine alla *path analysis*, in quanto permette ai ricercatori di testare

le relazioni ipotizzate tra i costrutti (Weston & Gore, 2006). Una differenza tra il SEM e gli altri metodi, e un vantaggio del SEM, è la sua capacità di stimare e testare le relazioni tra i costrutti. Rispetto ad altri modelli lineari generali, in cui i costrutti possono essere rappresentati con una sola misura e l'errore di misura non viene modellato, il SEM consente l'uso di misure multiple per rappresentare i costrutti e affronta il problema dell'errore specifico della misura. Questa differenza è importante in quanto consente ai ricercatori di stabilire la validità di costrutto dei fattori (Weston & Gore, 2006): la sua capacità di modellare ipotesi complesse e di tenere conto dell'errore di misurazione crea un metodo notevole e flessibile per analizzare i dati raccolti nel tempo (Newsom, 2015). Pensare al modello di equazioni strutturali come a una combinazione di analisi fattoriale e di *path analysis* permette di concettualizzare le due componenti principali del SEM: il modello di misurazione e il modello strutturale. Il modello di misurazione descrive le relazioni tra le variabili osservate (ad es., gli strumenti) e il costrutto (o i costrutti) che si ipotizza misurino tali variabili, mentre il modello strutturale descrive le interrelazioni tra i costrutti (Weston & Gore, 2006). L'analisi fattoriale confermativa è fondamentale per i modelli di equazioni strutturali, dato che rappresenta la modalità di verifica del modello di misurazione sopracitato (Weston & Gore, 2006; Wang & Wang, 2012).

L'analisi fattoriale confermativa è utilizzata in situazioni in cui le relazioni tra variabili oggetto di studio sono in parte note, per conoscenza teorica o risultati empirici (Wang & Wang, 2012). In questo caso, lo scopo dei ricercatori è quello di determinare e confermare se strutture fattoriali delle scale nello strumento sotto indagine siano quelle ipotizzate. Nell'applicazione della CFA, i ricercatori sono interessati soprattutto a valutare in che misura un insieme di indicatori o item di un particolare strumento misurino effettivamente le variabili o fattori latenti che sono stati progettati per misurare (Wang &

Wang, 2012). A differenza dell'analisi fattoriale esplorativa, nell'analisi fattoriale confermativa le relazioni tra fattori possono essere specificate in modo flessibile su una base teorica o sulla base di risultati empirici. Inoltre, nell'analisi fattoriale confermativa gli indicatori osservati o item saturano solo sui fattori che sono ipotizzati misurare: per questo motivo, un modello di analisi fattoriale confermativa non è solamente più significativo dal punto di vista teorico, ma anche più parsimonioso, dato che le saturazioni degli indicatori sui fattori irrilevanti sono tutti fissati a priori a zero, riducendo il numero di parametri da stimare (Wang & Wang, 2012). Inoltre, la CFA permette la correlazione degli errori di misurazione, mentre ciò non è previsto per un'analisi fattoriale esplorativa. Infine, l'analisi fattoriale confermativa può essere condotta simultaneamente in più gruppi: questo permette di testare l'invarianza della misura tra gruppi. Nella CFA, il legame tra gli indicatori osservati o gli item e i fattori è rappresentato dalla saturazione dei fattori, che sono i percorsi di regressione dai fattori ai corrispondenti indicatori osservati. Il coefficiente angolare (o pendenza) della regressione di un indicatore osservato su un fattore è la saturazione dell'indicatore su quel fattore e il termine residuo associato è il corrispondente errore di misura dell'indicatore. Pertanto, la misura di un indicatore osservato è suddivisa in errore di misurazione e misura del fattore sottostante. Di conseguenza, nel modellare le relazioni tra fattore e variabili latenti, le relazioni stimate tra le variabili latenti saranno prive degli effetti degli errori di misura (Wang & Wang, 2012).

2.4.1 L'Analisi Fattoriale Confermativa di primo e di secondo ordine

Nel caso di CFA con fattori multipli, la struttura di varianza/covarianza dei fattori può essere analizzata ulteriormente introducendo fattori di secondo ordine nel modello. Questo è possibile se i fattori di primo ordine sono correlati in modo robusto tra loro e se

è possibile ipotizzare che i fattori di secondo ordine spieghino la variabilità dei fattori di primo ordine (Chen et al., 2005; Wang & Wang, 2012). Per esempio, per testare l'esistenza di un fattore generale d'intelligenza che sta alla base di un'ampia gamma di specifiche abilità relative all'intelligenza, è possibile ipotizzare che le abilità specifiche (ognuna valutata da più item) siano fattori di ordine inferiore, e l'intelligenza generale sia il fattore di ordine maggiore, che spiega la comunanza tra le abilità specifiche (Chen et al., 2005).

Chen e collaboratori (2005) affermano che un modello fattoriale di secondo ordine presenta numerosi potenziali vantaggi rispetto al modello fattoriale di primo ordine. In primo luogo, il modello di secondo ordine può verificare se il fattore di ordine superiore ipotizzato rende effettivamente conto dello schema di relazioni tra i fattori del primo ordine. In secondo luogo, un modello del secondo ordine dà una struttura al modello di covarianza tra i fattori del primo ordine, spiegando la covarianza in modo più parsimonioso e con meno parametri (Gustafsson & Balke, 1993; Rindskopf & Rose, 1988). In terzo luogo, un modello del secondo ordine separa la varianza dovuta a fattori specifici dall'errore di misurazione, portando a una stima teoricamente priva di errori dei fattori specifici. La varianza unica di ciascun fattore del primo ordine che non è condivisa dal fattore comune di secondo ordine rappresenta i fattori specifici. Questi fattori specifici sono rappresentati dal disturbo di ciascun fattore del primo ordine. Infine, i modelli di secondo ordine possono fornire una utile semplificazione nell'interpretazione di strutture di misura complesse, come i modelli multi-tratto multi-metodo (Chen et al., 2005).

Chen e collaboratori (2005) individuano aspetti di cui è necessario tenere conto nel testare l'invarianza di misura dei modelli di secondo ordine. In primo luogo, l'invarianza delle saturazioni deve essere testata sia per i fattori del primo ordine che per quelli del

secondo ordine. In secondo luogo, l'invarianza dell'intercetta deve essere verificata sia per le variabili misurate sia per i fattori del primo ordine. Le medie dei fattori del primo ordine sono funzione delle intercette delle variabili misurate e delle saturazioni e delle medie dei fattori del primo ordine; la media dei fattori del secondo ordine è funzione delle intercette dei fattori del primo ordine e dei caricamenti e delle medie dei fattori del secondo ordine. Infine, oltre a verificare l'invarianza della varianza residua delle variabili osservate, è necessario verificare anche l'invarianza delle misure di disturbo (fattori specifici) dei fattori del primo ordine. Quando si raggiunge questo livello di invarianza, significa che le condizioni di disturbo (fattori specifici, ad es., l'abilità spaziale nel caso dell'intelligenza) dei fattori di ordine inferiore sono equivalenti tra i gruppi (Chen et al., 2005).

2.4.2 Gli indici di *fit* nell'Analisi Fattoriale Confermativa

Una volta stimato il modello fattoriale confermativo, è necessario stimare la *fit* o adattamento del modello ai dati: l'obiettivo è quello di determinare se le associazioni tra le variabili misurate e latenti nel modello stimato dal ricercatore riflettono adeguatamente le associazioni osservate nei dati (Weston & Gore, 2006). Gli statistici concordano sul fatto che i ricercatori dovrebbero valutare l'idoneità in termini di (a) significatività e forza dei parametri stimati, (b) varianza rappresentata dalle variabili endogene osservate e latenti, e (c) quanto il modello complessivo si adatta ai dati osservati, come indicato da una serie di indici di idoneità. Sebbene affrontare i primi due criteri sia un compito abbastanza semplice, esiste un notevole disaccordo su quali siano i valori accettabili per gli indici di adattamento globali, ed esistono numerosi indici per la valutazione della *fit* del modello (Weston & Gore, 2006). Il concetto più rigoroso di *fit* suggerisce come il

modello debba replicare esattamente i dati osservati, mentre una seconda prospettiva afferma che i modelli che approssimano i dati osservati siano accettabili (Quintana & Maxwell, 1999). Hoyle & Panter (1995) raccomandano di riportare numerosi indici di *fit* del modello. Di seguito, i principali indici riportati.

GFI (Goodness of Fit Index) e χ^2 . Essi rappresentano indici di *fit* assoluta, che valutano direttamente la misura in cui un modello si adatta ai dati osservati. Questi indici sono utili per confrontare i modelli quando si testano ipotesi concorrenti. Il GFI è analogo a R^2 , utilizzato nella regressione per riassumere la varianza spiegata in una variabile dipendente, ma il GFI si riferisce alla varianza rappresentata dall'intero modello. Tuttavia, i ricercatori riportano il GFI con meno frequenza di χ^2 (Weston & Gore, 2006). Dato che il valore χ^2 rappresenta un test di mis-specificazione del modello, un χ^2 significativo suggerirà che il modello non si adatta ai dati del campione. Al contrario, un χ^2 non significativo indicherà un modello che si adatta bene ai dati. Sebbene l'indice di adattamento assoluto più comunemente riportato sia il χ^2 , questa statistica presenta due limitazioni. In primo luogo, essa verifica se il modello risulta essere un *fit* perfetto per i dati, raro da verificare. In secondo luogo, come per la maggior parte delle statistiche, le grandi dimensioni del campione aumentano la potenza, con il risultato di una significatività con piccole dimensioni dell'effetto (Weston & Gore, 2006). Di conseguenza, un χ^2 non significativo può essere improbabile, anche se il modello può adattarsi perfettamente ai dati osservati. Nonostante queste limitazioni, i ricercatori riportano il χ^2 quasi universalmente (Martens, 2005) e di solito considerano altri indici per determinare se il *fit* del modello è accettabile.

CFI (Comparative Fit Index). Il CFI (Bentler, 1990) è un esempio di indice di *fit* incrementale: questo tipo di indice confronta il miglioramento del *fit* del modello del

ricercatore rispetto a un modello più ristretto, chiamato modello di indipendenza o modello nullo, che non specifica alcuna relazione tra le variabili. Il CFI varia da 0 a 1.0, con valori più vicini a 1.0 che indicano un migliore adattamento.

RMSEA (Root Mean Squared Error of Approximation). L'indice RMSEA (Steiger, 1990; Steiger & Lind, 1980) corregge per la complessità del modello. Di conseguenza, quando due modelli spiegano ugualmente bene i dati osservati, il modello più semplice avrà un valore RMSEA più favorevole. Un valore RMSEA di .00 indica che il modello si adatta esattamente ai dati. Una pratica adottata dai ricercatori consiste nel fornire anche l'intervallo di confidenza del 90% associato (90% CI), che incorpora l'errore di campionamento associato all'RMSEA stimato.

SRMR (Standardized Root Mean Square Residual). L'indice SRMR (Bentler, 1995) si basa sui residui di covarianza, con valori più piccoli che indicano un migliore *fit*. L'SRMR è una sintesi della differenza esistente tra i dati osservati e il modello. L'SRMR è la media assoluta di tutte le differenze tra le correlazioni osservate e quelle implicite del modello. Una media pari a zero indica l'assenza di differenze tra i dati osservati e le correlazioni implicite nel modello; pertanto, un SRMR pari a .00 indica un adattamento perfetto.

Inizialmente, le linee guida per un *fit* accettabile includevano un χ^2 non significativo, CFI maggiore di .90, RMSEA minore di .10 con un limite superiore massimo dell'IC del 90% pari a .10 e SRMR minore di .10 (Weston & Gore, 2006). I parametri per l'accettabilità della *fit* del modello sono però dibattuti: Hu & Bentler (1998) suggeriscono un *cutoff* minimo di .95 per CFI e un massimo *cutoff* di .06 per RMSEA. I ricercatori sostengono inoltre come fattori di numerosità campionaria, complessità del modello e grado di misspecificazione possano influenzare valori di *cutoff* appropriati:

l'applicazione inappropriata dei nuovi criteri di *cutoff* potrebbe portare al rifiuto errato di modelli accettabili quando le dimensioni del campione sono inferiori a $n = 500$ e quando i modelli non sono complessi. D'altro canto, l'utilizzo dei vecchi criteri per valutare modelli complessi stimati con campioni più grandi di $n = 500$ può portare all'accettazione errata di modelli mal specificati. La ricerca empirica suggerisce che gli indici di *fit* che non soddisfano le precedenti linee guida (i.e., $CFI \geq .90$, $RMSEA \leq .10$, e $SRMR \leq .10$) probabilmente non sarebbero accettabili e indica che i modelli con indici che superano i nuovi criteri sarebbero accettabili (i.e., $CFI \geq .95$, $RMSEA \leq .06$, e $SRMR \leq .08$). Ciò suggerisce che quando i valori di CFI sono compresi tra .90 e .95, i valori di RMSEA tra .05 e .10, e SRMR tra .08 e .15 i ricercatori dovrebbero considerare la dimensione del campione utilizzato per stimare il modello, utilizzando criteri più severi per campioni maggiori di $n = 500$, e la complessità del modello, utilizzando criteri più severi per modelli meno complessi. (Weston & Gore, 2006).

Nell'ambito dell'invarianza, tali indici sono tuttavia utili esclusivamente per valutare l'adattamento del modello configurale. Quando si confrontano modelli nidificati con vincoli di uguaglianza tra i gruppi (ad es., si pongono tutte le saturazioni uguali tra i gruppi e si confronta il modello di invarianza metrica con quello configurale), oltre al test del chi quadrato per la differenza tra modelli (*test of close fit*), è opportuno valutare la differenza tra indici di *fit* assoluti e comparativi che dipendano meno dalla numerosità campionaria.

A tal proposito, Chen (2007) suggerisce per un campione di meno di 300 soggetti, quando la dimensione del campione è diseguale e il modello di non invarianza è uniforme, i seguenti criteri di cutoff: per testare l'invarianza metrica, una variazione di .005 in CFI, integrata da una variazione di .010 in RMSEA o una variazione di .025 in SRMR

indicherebbe la non invarianza; per testare l'invarianza scalare e stretta, una variazione di .005 in CFI, integrata da una variazione di .010 in RMSEA o una variazione di .005 in SRMR indicherebbe la non invarianza. Chen (2007) suggerisce inoltre i seguenti criteri di cutoff più rigorosi per un campione superiore a 300 soggetti, con dimensione del campione uniforme tra i gruppi: per verificare l'invarianza metrica, una variazione di .010 nel CFI, integrata da una variazione di .015 nell'RMSEA o da una variazione di .030 nell'SRMR, indicherebbe la non invarianza; per verificare l'invarianza scalare e stretta, una variazione di .010 nel CFI, integrata da una variazione di .015 nell'RMSEA o da una variazione di .010 nell'SRMR indicherebbe la non invarianza. Tra i tre indici, il CFI è stato scelto come criterio principale perché RMSEA e SRMR tendono a rifiutare eccessivamente un modello invariante quando la dimensione del campione è piccola (Chen, 2007).

3. STUDIO EMPIRICO

3.1 Obiettivi e Ipotesi di ricerca

Come evidenziato nei precedenti capitoli, l'interesse per il costrutto di chiamata professionale è in crescente aumento, e le ricerche condotte risultano essere in esponenziale crescita, in tutte le culture (Vianello et al., 2018). Nonostante questa crescente attenzione per il campo di studi, rimane tutt'ora aperta la questione riguardante l'equivalenza nella concettualizzazione del costrutto di chiamata tra culture, e la generalizzabilità dei risultati di tali ricerche risulta essere messa in discussione. Come affermato dai ricercatori Hui e Triandis (1985), risulta fondamentale in ogni ricerca cross-culturale preoccuparsi di garantire l'equivalenza (i.e. la comparabilità) durante la verifica delle differenze cross-culturali.

La seguente ricerca empirica ha lo scopo di investigare la generalizzabilità cross-culturale della concettualizzazione di chiamata tra sette campioni appartenenti a sei nazioni (India, Turchia, Cina, Italia, Stati Uniti e Paesi Bassi), testandone l'invarianza della misura. A questo scopo, è stata condotta un'analisi fattoriale confermativa di secondo ordine per singoli paesi e, successivamente, un'analisi multi-gruppo dell'invarianza della misura.

3.2 Metodo

3.2.1 L'approccio Open Science

Nel Novembre 2021, l'Organizzazione delle Nazioni Unite per l'Educazione, la Scienza e la Cultura (UNESCO) ha adottato “*UNESCO Recommendation on Open Science*”, definendo la scienza aperta come "un costrutto inclusivo che combina vari

movimenti e pratiche che mirano a rendere la conoscenza scientifica multilingue apertamente disponibile, accessibile e riutilizzabile per tutti, ad aumentare le collaborazioni scientifiche e la condivisione delle informazioni a beneficio della scienza e della società, e ad aprire i processi di creazione, valutazione e comunicazione della conoscenza scientifica agli attori della società al di là della comunità scientifica tradizionale" (UNESCO, 2021).

In particolare, il campo della psicologia si trova al centro di un periodo di autoesame e auto-miglioramento (Vazire, 2018). La cosiddetta “rivoluzione della credibilità” (Angrist & Pischke, 2010), o crisi della replicabilità, ha portato numerosi cambiamenti in merito agli standard con i quali la scienza psicologica viene valutata (Vazire, 2018). Due dei principali temi e proposte per il miglioramento della disciplina sono quello della trasparenza e della preregistrazione (Vazire, 2018). Il primo si riferisce alla necessità di una maggiore trasparenza nel processo di ricerca e di pubblicazione (Nosek et al., 2015): questo include una migliore documentazione e accessibilità alle procedure, ai materiali, ai dati e alle analisi, richiedendo ai ricercatori un report completo di tutti i risultati, impedendo un report selettivo (Vazire, 2018). La seconda proposta di miglioramento riguarda la preregistrazione dei nuovi studi: essa consiste nel mettere per iscritto il proprio metodo, piano di analisi e i risultati previsti prima di raccogliere dati e poi rendere questo piano pubblicamente disponibile. Questo permette di individuare quali decisioni di analisi e quali risultati sono stati predetti a priori e possono essere interpretati come confermativi.

Questa ricerca cross-culturale si serve dell’Open Science Framework (OSF) per permettere un approccio di Scienza Aperta e riproducibile.

3.2.2 Partecipanti

Hanno preso parte alla ricerca 2,490 studenti universitari appartenenti a sei paesi (Turchia, India, Cina, Italia, Stati Uniti e Paesi Bassi) che hanno compilato il questionario in sei lingue (Turco, Hindi, Cinese, Italiano, Inglese, Olandese), per un totale di sette campioni (Turchia, India-Hindi, India-Inglese, Cina, Italia, Stati Uniti e Paesi Bassi). I partecipanti hanno compilato il questionario nella lingua ufficiale della propria nazione, e i soggetti indiani hanno potuto scegliere se compilare il questionario in lingua hindi o inglese, dato che entrambe le lingue risultano ufficiali nello stato dell'India. Le statistiche descrittive sono state calcolate con il software Jasp. Le Tabelle 3.1 e 3.2 riportano le caratteristiche di età e sesso dei partecipanti divise per nazioni.

I campioni risultano piuttosto bilanciati per le statistiche descrittive di sesso ed età. Fanno eccezione una larga percentuale di partecipanti uomini per i paesi India e Paesi Bassi (India-HI 60.2%, India-EN 68.9%, Paesi Bassi 61%) e di partecipanti donne nel sottocampione cinese (74%). I partecipanti presentano un'età media compresa tra i 19 anni per i Paesi Bassi ($M=19$, $SD=1.45$) e la Cina ($M=19.92$, $SD=3.37$) e i 24 anni per il sottocampione Indiano, sia in lingua hindi che inglese (India-HI $M=24.53$, $SD=4.52$; India-EN $M=24.89$, $SD=3.33$). Anche la numerosità campionaria tra i sottocampioni risulta piuttosto omogenea ($M=355.71$, $SD=66.57$): il sottocampione India-HI risulta essere il meno numeroso ($n=252$), mentre quello turco il più numeroso ($n=486$).

I partecipanti sono tutti studenti universitari e appartengono esclusivamente a corsi Triennali (i.e., Bachelor's degree) o Magistrali (i.e., Master's degree) di studi di business, economia o studi amministrativi.

Età	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Min	18	19	20	17	18	18	17
Max	36	50	42	60	58	50	32
<i>M</i>	21.89	24.53	24.89	19.92	22.05	22.44	19
<i>SD</i>	1.85	4.51	3.33	3.37	3.72	3.93	1.45

Tabella 3.1, Età dei partecipanti

Sesso	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Tot campione	486	252	393	317	352	354	336
Uomini	265 (54.9%)	133 (60.2%)	253 (68.9%)	50 (26%)	153 (51%)	140 (39.5%)	205 (61%)
Donne	216 (44.7%)	87 (39.4%)	111 (30.2%)	145 (74%)	147 (49%)	213 (60.2%)	129 (39%)

Tabella 3.2, Sesso dei partecipanti

Sono state raccolte ulteriori statistiche descrittive, riportate suddivise per paesi: la Tabella 3.3 è relativa all'etnia, la Tabella 3.4 alla religione e al grado in cui essa viene praticata, la Tabella 3.5 indica l'anno relativo al corso di studi e la Tabella 3.6 è relativa all'esperienza lavorativa. Per quanto riguarda l'etnia, l'India e la Cina risultano essere i campioni più etnicamente omogenei (India-HI 84.1% asiatico del sud, 14.9% Altro; India-EN 93.7% asiatico del sud, 5.4% Altro; Cina 95.5% asiatico confuciano, 3.6% Altro), Turchia, Italia e Paesi Bassi campioni più variabili (Turchia 48.6% caucasico, 27.9% mediorientale, 14.9% altro; Italia 38.2% caucasico, 47.3% latino europeo, 11.9% altro; Paesi Bassi 48.8% caucasico, 4.5% nordafricano, 33.3% altro), mentre il campione Statunitense risulta il più eterogeneo (21.6% caucasico, 24.9% ispanico o latino spagnolo, 20.1% asiatico del sud, 17.4% asiatico confuciano, 4.2% Altro). Per quanto riguarda la religione, il campione turco risulta essere per la maggior parte musulmano sunnita (64.2%), il campione indiano in maggioranza indu (India-HI 71.4%; India-EN 70.5%), il campione cinese per il 42.9% ateo e per il 47.7% altro/nessuna religione in particolare. Per il campione italiano la religione più praticata è il cattolicesimo (39.8%), mentre il 25.1% del campione si dichiara ateo e il 15.1% agnostico. Il campione statunitense risulta più vario, dichiarandosi per il 22.9% cattolico romano, per il 14.4% musulmano, per il 5.4% ebraico, per il 9% indu, 6.2% agnostico e 5.6% ateo, mentre il 26.8% del campione non presenta nessuna religione in particolare. Il campione relativo ai Paesi Bassi è composto per il 21.2% da atei, 7.5% protestati, 8.4% cattolici romani e 41.1% individui che preferiscono non rispondere. Per tutti i campioni, tranne quello Indiano, i partecipanti frequentano maggiormente un corso *Undergraduate*. I campioni statunitense e dei Paesi Bassi hanno più anni di esperienza lavorativa ($M=3.36$; $M=3.20$), mentre il campione cinese presenta il minor numero di anni di esperienza media (0.17%).

Etnia	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Nativo Americano o Nativo dell'Alaska	0	0	1 (0.3%)	0	0	3 (0.9%)	0
Nero o Afroamericano	3 (0.6%)	0	0	0	0	19 (5.7%)	3 (0.9%)
Caucasico	234 (48.6%)	1 (0.5%)	1 (0.3%)	0	113 (38.2%)	72 (21.6%)	162 (48.8%)
Asiatico Confuciano	2 (0.4%)	0	1 (0.3%)	187 (95.5%)	1 (0.3%)	58 (17.4%)	7 (2.1%)
Ispanico o latino spagnolo	3 (0.6%)	0	0	0	4 (1.4%)	83 (24.9%)	1 (0.3%)
Latino europeo	1 (0.2%)	0	0	0	140 (47.3%)	0	11 (3.3%)
Mediorientale	139 (27.9%)	0	0	1 (0.5%)	1 (0.3%)	16 (4.8%)	11 (3.3%)
Nativo hawaiano o altro isolano del Pacifico	0	0	0	0	1 (0.3%)	1 (0.3%)	0
Nordafriicano	11 (2.3%)	0	0	0	1 (0.3%)	0	15 (4.5%)
Asiatico del sud	11 (2.3%)	185 (84.1%)	343 (93.7%)	0	0	67 (20.1%)	10 (3%)
Africano sub- sahariano	5 (1%)	1 (0.5%)	0	0	0	0	3 (0.9%)
Altro/Mixed/Preferisc o non rispondere	72 (14.9%)	33 (14.9%)	20 (5.4%)	7 (3.6%)	35 (11.9%)	14 (4.2%)	109 (33.8%)

Tabella 3.3, Etnia dei partecipanti

Religione	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Agnostico	23 (4.8%)	2 (0.9%)	17 (4.6%)	9 (4.6%)	45 (15.1%)	22 (6.2%)	20 (6%)
Ateo	15 (3.1%)	12 (5.5%)	14 (3.8%)	84 (42.9%)	75 (25.1%)	20 (5.6%)	71 (21.3%)
Buddista	0	1 (0.5%)	2 (0.5%)	9 (4.6%)	2 (0.7%)	12 (3.4%)	3 (0.9%)
Indù	0	157 (71.4%)	258 (70.5%)	0	0	32 (9%)	3 (0.9%)
Ebraico	3 (0.6%)	0	3 (0.8%)	0	0	19 (5.4%)	3 (0.9%)
Musulmano	Sunnita: 310 (64.2%) Sciita: 18 (3.7%)	2 (0.9%)	7 (1.9%)	1 (0.5%)	1 (0.3%)	51 (14.4%)	39 (11.7%)
Ortodosso	9 (1.9%)	0	0	0	0	13 (3.7%)	4 (1.2%)
Protestante	0	21 (9.5%)	9 (2.5%)	0	1 (0.3%)	9 (2.5%)	25 (7.5%)
Cattolico Romano	3 (0.6%)	2 (0.9%)	15 (4.1%)	0	119 (39.8%)	81 (22.9%)	28 (8.4%)
Altro/Preferisco non rispondere/Nessuna in particolare	102 (21%)	25 (10.4%)	41 (11.2%)	93 (47.4%)	54 (18.7%)	95 (26.8%)	138 (41.1%)
Praticante Attivo	227 (47.4%)	138 (63%)	167 (45.8%)	63 (32.3%)	64 (22.4%)	155 (43.7%)	66 (19.8%)
Non Praticante	252 (52.6%)	81 (36.9%)	198 (54.2%)	132 (67.6%)	222 (77.6%)	199 (56.2%)	268 (80.2%)

Tabella 3.4, Religione dei partecipanti e pratica attiva

Tabella 3.5, Anno di studio dei partecipanti

Anno di studi	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Primo anno Laurea Triennale	5 (1%)	70 (32%)	5 (1.4%)	69 (35.2%)	50 (16.7%)	11 (3.1%)	303 (90.7%)
Secondo anno Laurea Triennale	184 (38.1%)	5 (2.3%)	1 (0.3%)	80 (40.8%)	68 (22.7%)	54 (15.3%)	24 (7.2%)
Terzo anno Laurea Triennale	168 (34.8%)	17 (7.8%)	0	34 (17.3%)	91 (30.3%)	123 (34.7%)	1 (0.3%)
Quarto anno Laurea Triennale	120 (24.8%)	0	0	6 (3.1%)	0	32 (9%)	0
Primo anno Laurea Magistrale	2 (0.4%)	95 (43.4%)	331 (90.4%)	0	50 (16.7%)	4 (1.1%)	0
Secondo anno Laurea Magistrale	1 (0.2%)	16 (7.3%)	25 (6.8%)	7 (3.6%)	37 (2.3%)	37 (12.3%)	0
Terzo anno Laurea Magistrale	2 (0.4%)	16 (7.3%)	1 (0.3%)	0	1 (0,3%)	94 (26.6%)	2 (0.6%)
Altro/Preferisco non rispondere	1 (0.2%)	0	0	0	3 (1%)	36 (10.2%)	0

Anni di sperienza lavorativa	Turchia	India HI	India EN	Cina	Italia	Stati Uniti	Paesi Bassi
Min	0	0	0	0	0	0	0
Max	11	11	11	11	11	11	11
<i>M</i>	1.26	1.75	2.51	0.17	1.46	3.36	3.20
<i>SD</i>	1.75	2.38	2.52	0.86	1.83	2.78	1.76

Tabella 3.6, Anni di esperienza lavorativa dei partecipanti

3.2.3 Procedura

Lo studio è stato condotto mediante la somministrazione di un questionario self-report a tutti i partecipanti in modalità da remoto, dopo aver letto e approvato il modulo di consenso informato. Il campionamento si è svolto con una modalità di comodo, e i partecipanti volontari hanno compilato il questionario sulla piattaforma Qualtrics. I partecipanti dei paesi India, Turchia, Stati Uniti e Paesi Bassi hanno ricevuto un incentivo accademico per la partecipazione. Gli studenti italiani sono stati reclutati attraverso social media o con un campionamento a catena.

Il questionario somministrato era di tipo autovalutativo, e ai partecipanti è stato richiesto di indicare il proprio grado di accordo o disaccordo su una serie di item che compongono le scale per la misura dei costrutti d'interesse, sulla base delle proprie opinioni e preferenze.

3.2.4 Misure e Strumenti

Il questionario presentato ai partecipanti include misure non utilizzate in questo studio. In particolare, sono stati raccolti dati in merito ai costrutti di *work values* (36 items), collettivismo (6 items), *work centrality* (5 items), percezione della chiamata (2 items), motivazione alla carriera (3 items), ansia (4 items), ottimismo (7 items), *work volition* (4 items), *perceived employability* (3 items), soddisfazione nella vita (5 items), significato nella vita (5 items) e costruzione e sviluppo di carriera (7 items).

La scala utilizzata per la valutazione della chiamata professionale è la *Unified Multidimensional Calling Scale* (UMCS) nella sua versione a 28 items, trattata nel

Paragrafo 1.3.1. Di seguito, in Tabella 3.7, la scala a 28 items riportata con la nomenclatura degli items.

UMCS1	Pass_1	Nutro per ciò che studio una vera passione
UMCS2	Pass_2	Mi piace quello che studio più di ogni altra cosa
UMCS3	Pass_3	Questi studi mi danno un'enorme soddisfazione
UMCS4	Pass_4	Ciò che studio mi fa vivere esperienze intense ed emozionanti
UMCS5	Sac_1	Continuerei questi studi anche di fronte a grandi ostacoli
UMCS6	Sac_2a	Posso affrontare molte rinunce pur di continuare questi studi
UMCS7	Sac_3	Sono disposto a grandi sacrifici pur di continuare questi studi
UMCS8	Sac_4	Sono pronto ad affrontare molte difficoltà pur di continuare questi studi
UMCS9	TrS_1	Sto seguendo questo percorso perché credo di essere stato "chiamato" a farlo
UMCS10	TrS_2b	Sono stato "chiamato", a seguire questo percorso da qualcosa oltre me
UMCS11	TrS_3	Credo di essere stato "chiamato" a seguire questo percorso
UMCS12	TrS_4	Mi sento "chiamato"; verso questo percorso da qualcosa più grande di me
UMCS13	Pro_1	Un aspetto importante del mio percorso accademico e professionale è che contribuisca a soddisfare i bisogni della società
UMCS14	Pro_2	Cerco sempre di valutare quanto il mio percorso accademico e professionale potrà essere di beneficio per la società
UMCS15	Pro_3	È importante per me che il mio percorso accademico e professionale sia utile in qualche modo alla società
UMCS16	Pro_4	È importante per me che il mio percorso accademico e professionale contribuisca a rendere il mondo un posto migliore
UMCS17	Per_1	Anche quando non studio penso spesso alle materie del mio corso
UMCS18	Per_2	Ciò che studio è sempre nei miei pensieri
UMCS19	Per_3	Le mie giornate avrebbero meno senso se non pensassi spesso a ciò che studio
UMCS19A	Per_5	Non posso fare a meno di pensare a ciò che studio
UMCS20	Per_4	Penso a ciò che studio in ogni momento della giornata

UMCS21	Pur_1	Vedo la mia carriera accademica e professionale come un percorso per realizzarmi nella vita
UMCS22	Pur_2	Il mio percorso accademico e professionale è importante per dare significato alla mia vita
UMCS23	Pur_3	Il mio percorso accademico e professionale mi aiuta a realizzare lo scopo della mia vita
UMCS24	Pur_4	Il mio percorso accademico e professionale dà senso alla mia vita
UMCS25	Ide_1	Quello che studio sarà sempre parte della mia vita
UMCS26	Ide_2	Quello che studio sarà sempre parte di me
UMCS27	Ide_3	Quello che studio è parte del mio destino
UMCS27A	Ide_5	Quello che studio è parte della mia identità
UMCS28	Ide_4	Quello che studio è parte del mio futuro

Tabella 3.7 *Unified Multidimensional Calling Scale* (UMCS28)

La *Unified Multidimensional Calling Scale* misura sette *facets*: passione (*Pass*), sacrificio (*Sac*), evocazione trascendente (*TrS*), orientamento prosociale (*Pro*), pervasività (*Per*), significato (*Pur*), identità (*Ide*). Queste sette dimensioni identificate sintetizzano le due prospettive del costrutto di chiamata (neoclassica e moderna), possiede buone proprietà psicometriche ed è invariante a livello stretto (varianze d'errore equivalenti) attraverso il tempo e i domini di studio (Vianello et al., 2018).

I partecipanti hanno risposto tramite una scala Likert a cinque punti: 1 – fortemente in disaccordo, 2 – parzialmente in disaccordo, 3 – né d'accordo né in disaccordo, 4 – parzialmente d'accordo, 5 – fortemente d'accordo.

La scala è stata inizialmente sviluppata e validata in italiano (Vianello et al., 2018), e successivamente tradotta in inglese. La versione inglese è stata poi tradotta nella lingua

ufficiale di ogni paese partecipante allo studio e ritradotta in inglese da un diverso ricercatore per valutare la validità delle versioni in lingua.

3.3 Analisi dei dati

Le analisi descrittive del campione discusse nel paragrafo 3.2.2 “Partecipanti” sono state condotte attraverso il software Jasp.

Le analisi effettuate sui dati raccolti relativi alla *Unified Multidimensional Calling Scale* sono state condotte allo scopo di testare l’invarianza del costrutto di chiamata tra le culture coinvolte. È stata effettuata in primis un’analisi della struttura fattoriale della scala per ogni paese, mediante fattoriale confermativa di secondo ordine (CFA). Dopo aver verificato che il modello utilizzato fosse una buona rappresentazione dei dati raccolti in tutti i paesi coinvolti, è stata condotta un’analisi dell’invarianza della misura multi-gruppo tra i paesi coinvolti. Tutte le analisi dei dati sono effettuate con il software Mplus 7 (Muthén & Muthén, 1998-2012).

3.4 Risultati

3.4.1 Risultati dell’analisi fattoriale confermativa di secondo ordine

È stata effettuata un’analisi fattoriale confermativa per ogni paese. Ogni CFA include un fattore di secondo ordine, la chiamata, e sette fattori di primo ordine, corrispondenti ai sette *facets* della chiamata (passione, sacrificio, evocazione trascendente, orientamento prosociale, pervasività, significato, identità). È ipotizzato che il fattore di secondo ordine influenzi i punteggi latenti dei sette fattori di primo ordine,

ognuno dei quali rappresenta la varianza di quattro indicatori (ad esempio, gli items Pro_1, Pro_2, Pro_3 e Pro_4 rappresentati dal fattore orientamento prosociale). Ogni item carica quindi sul fattore di primo ordine corrispondente, e i sette fattori di primo ordine caricano sul singolo fattore di secondo ordine che rappresenta il costrutto di chiamata.

Per la valutazione della *fit* del modello in ogni paese, sono utilizzati gli indici di *fit*. In particolare, per il *comparative fit index* (CFI) è ritenuto accettabile un valore superiore a .90 (Hu & Bentler, 1995). Per l'indice RMSEA (*root mean square of approximation*) valori inferiori a .08 indicano una *fit* accettabile, e valori inferiori a .06 una buona *fit* (Hu & Bentler, 1999). Per l'indice SRMR (*standardized root mean square of residuals*) valori inferiori o uguali a .08 sono interpretati come indicatore di una buona *fit* (Hu & Bentler, 1999). La Tabella 3.8 presenta i risultati delle analisi fattoriali confermatrice per paese.

I valori degli indici di *fit* risultano buoni per tutti i gruppi tranne che per il campione India-HI, che presenta $CFI < .90$, $RMSEA > .06$ e $SRMR > .08$. È stato quindi rilevato un problema con la traduzione in Hindi dell'item UMCS11 o TrS_3 (“Credo di essere stato “chiamato” a seguire questo percorso”). In particolare, nella traduzione Hindi dell'item (“मैं यह मानता हूँ कि इस विषय में पढ़ाई के लिए आंतरिक प्रेरणा कम है।”), il termine usato per tradurre “chiamato” è stato probabilmente interpretato come una motivazione intrinseca o ispirazione interna invece che come una evocazione trascendente. Per questo motivo, l'item è stato rimosso dal modello, portando ad un miglioramento degli indici di *fit* del campione India-HI ($\Delta\chi^2 = 121.53$, $\Delta df = 26$, $p < .001$; $\Delta CFI = .03$; $\Delta RMSEA = .01$; $\Delta SRMR = .02$) e all'accettabilità degli indici di *fit*. È stata valutata anche la variazione del modello India-EN alla rimozione dell'item UMCS11, senza rilevare un miglioramento nella *fit*, dimostrando quindi il problema nella traduzione dell'item per India-HI. Il campione India-EN mantiene quindi l'item UMCS11.

Tabella 3.8, indici dell'analisi fattoriale confermativa per paesi

	N	n parametri	χ^2	df	χ^2/df	CFI	RMSEA	90% LL	90% UL	SRMR
India-EN	393	92	619.687	342	1.81195	0.941	0.045	0.04	0.051	0.054
<i>India-En senza item 11^a</i>		89	531.662	316		0.951	0.042	0.035	0.048	0.052
India-Hindi	252	92	666.492	342	1.948807	0.879	0.061	0.054	0.068	0.089
India-Hindi senza item 11 ^a	252	89	544.967	316	1.724579	0.91	0.054	0.046	0.061	0.068
Cina	317	92	585.549	342	1.712132	0.918	0.047	0.041	0.054	0.057
Turchia	486	92	638.861	342	1.868015	0.95	0.042	0.037	0.047	0.053
Italia	352	92	537.784	342	1.572468	0.953	0.04	0.034	0.047	0.056
Stati Uniti	354	92	647.665	342	1.893757	0.922	0.05	0.044	0.056	0.058
Paesi Bassi	336	92	635.48	342	1.858129	0.913	0.051	0.044	0.057	0.063

Nota. Le statistiche di *fit* per l'analisi fattoriale confermativa sono state valutate con i seguenti criteri: CFI $\geq$.90; RMSEA $\leq$.06; SRMR $\leq$.08 (Hu & Bentler, 1995; 1999).

^a Modello stimato senza l'item UMCS11 "Credo di essere stato "chiamato" a seguire questo percorso"

3.4.2 Risultati dell'analisi dell'invarianza della misura multi-gruppo

Una volta confermata la bontà della *fit* del modello, è stata testata l'invarianza della misura del costrutto di chiamata tra le culture prese in esame. L'invarianza della misura viene testata su quattro livelli, comparando una serie di modelli di equazioni strutturali multi-gruppo nidificati.

I livelli di invarianza testati, descritti nel dettaglio nel Paragrafo 2.3.1, sono l'invarianza della misura configurale (stesso item che satura sullo stesso fattore latente), metrica (saturazioni fattoriali di primo e secondo ordine equivalenti), scalare (stesse intercette degli items tra culture) e stretta (stesse varianze d'errore). Questi modelli sono nidificati: ad ogni successiva analisi vengono vincolati una serie di parametri ad essere uguali tra i paesi. Ogni livello di invarianza mantiene i vincoli precedenti, per cui il modello di invarianza stretta risulta essere il più parsimonioso, con il minor numero di parametri stimati. La Tabella 3.9 rappresenta le specificazioni del modello per verificare l'invarianza della misura. Nella Tabella, il primo cambiamento in una serie di vincoli per ogni tipo di parametro (ogni colonna) è indicato in corsivo.

Nella comparazione dei modelli nidificati, vengono adottate le soglie di accettabilità di *fit* suggerite da Chen (2007) introdotti nel Paragrafo 2.4.2. L'invarianza metrica è rifiutata con una modifica di $\leq .010$ in CFI, integrata ad una modifica di $\geq .015$ in RMSEA o di $\geq .030$ in SRMR; i modelli di invarianza scalare e stretta sono rifiutati con una modifica di $\leq .010$ in CFI, integrata ad una modifica di $\geq .015$ in RMSEA o di $\geq .010$ in SRMR (Chen, 2007).

Livello di invarianza	Loading del primo ordine	Loading del secondo ordine	Intercetta dell'indicatore	Medie del primo ordine	Medie del secondo ordine	Varianza dell'indicatore	Varianza del primo ordine	Varianza del secondo ordine
Configurale	Primo loading fissato a 1 in ogni fattore, tutti gli altri liberi	Loading di Passione fissato a 1, gli altri liberi	Libero in tutti i gruppi	Fissato a 0 in tutti i gruppi	Fissato a 0 in tutti i gruppi	Libero in tutti i gruppi	Libero in tutti i gruppi	Libero
Metrica primo ordine	<i>Primo loading fissato a 1 in ogni fattore, tutti gli altri fissati uguali tra i gruppi</i>	Loading di Passione fissato a 1, gli altri liberi	Libero in tutti i gruppi	Fissato a 0 in tutti i gruppi	Fissato a 0 in tutti i gruppi	Libero in tutti i gruppi	Libero in tutti i gruppi	Libero
Metrica secondo ordine	Primo loading fissato a 1 in ogni fattore, tutti gli altri fissati uguali tra i gruppi	<i>Loading di Passione fissato a 1, gli altri fissati uguali tra i gruppi</i>	Libero in tutti i gruppi	Fissato a 0 in tutti i gruppi	Fissato a 0 in tutti i gruppi	Libero in tutti i gruppi	Libero in tutti i gruppi	Libero

Scalare	Primo loading fissato a 1 in ogni fattore, tutti gli altri fissati uguali tra i gruppi	Loading di Passione fissato a 1, gli altri fissati uguali tra i gruppi	<i>Fissati uguali tra i gruppi</i>	<i>Fissato a 0 in Italia, libero negli altri gruppi</i>	Fissato a 0 in tutti i gruppi	Libero in tutti i gruppi	Libero in tutti i gruppi	Libero
Stretta	Primo loading fissato a 1 in ogni fattore, tutti gli altri fissati uguali tra i gruppi	Loading di Passione fissato a 1, gli altri fissati uguali tra i gruppi	Fissati uguali tra i gruppi	Fissato a 0 in Italia, libero negli altri gruppi	Fissato a 0 in tutti i gruppi	<i>Fissato uguale tra i gruppi</i>	<i>Fissato al valore precedente in Italia, libero negli altri gruppi</i>	<i>Fissato al valore precedente in Italia, libero negli altri gruppi</i>

Tabella 3.9 Modello di specificazione per il test dell'invarianza della misura.

Nota. Il primo cambiamento in una serie di vincoli per ogni tipo di parametro (colonna) è indicato in corsivo

Quando i vincoli di uguaglianza tra le nazioni hanno portato ad un aumento significativo del *misfit*, sono stati utilizzati indici di modifica per identificare gli specifici parametri che violavano i vincoli di invarianza di misura (Byrne et al., 1989). L'approccio è stato utilizzato per specificare modelli di invarianza parziale, in cui viene rilasciato il numero minimo di vincoli per ottenere una *fit* soddisfacente. Questo permette di ovviare all'obiettivo troppo rigido che i vincoli di invarianza debbano valere per tutti i parametri tra i gruppi (Byrne et al., 1989). Nell'invarianza parziale della misura, solo un sottoinsieme di parametri del modello è vincolato a essere invariante, mentre un altro sottoinsieme di parametri può variare tra i gruppi. Questo consente confronti appropriati tra i gruppi anche se non si ottiene l'invarianza completa della misura.

Gli indici di modifica stimati per i parametri vincolati sono calcolati utilizzando il metodo dei moltiplicatori di Lagrange; quindi, ci sono χ^2 differenze tra i modelli con un grado di libertà. Sono stati esaminati solo gli indici di modifica maggiori o uguali a 3.84, che riflettono il valore critico di $\chi^2_{(1)}$ con $p < .05$. Con le dimensioni del campione in esame ($N = 2.490$) e un livello alfa di .05, gli indici di modifica possono rilevare una differenza standardizzata di .06 nelle saturazioni e di .10 nelle intercette degli item. Gli indici di modifica superiori al valore critico indicano che il modello potrebbe essere significativamente migliorato rimuovendo il vincolo e liberando la stima dei parametri (Brown, 2006). Nel considerare gli indici di modifica, è stato verificato se la non invarianza fosse giustificata e se il parametro potesse essere interpretato in modo robusto, una volta liberato. In caso di assenza di una base sostanziale per liberare il vincolo sul parametro con l'indice di modifica più alto, è stato mantenuto il vincolo e considerata la possibilità di liberare il parametro con il secondo maggiore indice di modifica. Inoltre, per preservare l'interpretabilità dei punteggi delle sottoscale tra i paesi nel modello di

invarianza parziale, è stato mantenuto almeno il 50% dei vincoli di uguaglianza in ogni sottoscala della chiamata.

È stato quindi liberato un vincolo sui parametri alla volta, in un processo ciclico di esame dell'indice di modifica e di valutazione del modello di invarianza parziale della misura. L'adattamento del modello di invarianza parziale della misura è stato confrontato con l'adattamento del modello di invarianza della misura che testa il livello di invarianza precedente. Ad esempio, il modello scalare parziale è stato confrontato con il modello di invarianza metrica. La Tabella 3.10 riporta i risultati dei test di invarianza della misura.

Il modello di invarianza configurale presenta buoni indici di *fit*: le correlazioni tra i sette fattori di primo ordine possono essere significativamente spiegate da una struttura del secondo ordine con un fattore principale, il costrutto di chiamata. La struttura dei fattori del secondo ordine può quindi essere considerata uguale per tutti i gruppi.

L'invarianza delle saturazioni di fattori di primo e secondo ordine è stata testata utilizzando due modelli separati. Nel modello di specificazione, per l'invarianza metrica di primo ordine sono fissati ad essere uguali tra i paesi, mentre per quella di secondo ordine sono fissati anche i loadings dei fattori di primo ordine su quello di secondo ordine. Il vincolo di uguaglianza tra le saturazioni dei fattori non ha comportato una diminuzione della *fit* del modello: l'invarianza metrica può quindi essere accettata. Questo significa che gli item della scala UMCS sono ugualmente importanti nei vari paesi per definire le sette componenti della chiamata e che le sette componenti sono ugualmente importanti nei vari paesi per definire la chiamata.

L'invarianza della misura scalare lavora su vincoli a livello dell'intercetta, e il vincolo fissato su di esse ha diminuito il *fit* del modello: come è possibile osservare dalla

Tabella 3.10 Tabella 3.10, Indici di fit e comparazioni dei modelli di invarianza della misura

	n	χ^2	df	χ^2/df	CFI	RMSEA	LL	UL	SRMR	$\Delta\chi^2$	Δdf	ΔCFI	$\Delta RMSEA$	$\Delta SRMR$
Invarianza Configurale	643	4368.10	2395	1.82	0.929	0.048	0.046	0.05	0.064					
Invarianza Metrica primo ordine	518	4671.27	2520	1.85	0.922	0.049	0.047	0.051	0.074	303.17	125	-0.007	0.001	0.01
Invarianza Metrica secondo ordine	476	4833.40	2562	1.89	0.918	0.05	0.048	0.052	0.086	162.13	42	-0.004	0.001	0.012
Invarianza Scalare	351	6238.51	2687	2.32	0.872	0.061	0.059	0.063	0.098	1405.11	125	-0.046	0.011	0.012
Invarianza Scalare parziale	410	5192.927	2628	1.98	0.908	0.052	0.05	0.054	0.088	359.53	66	-0.01	0.002	0.002
Invarianza Stretta	235	6010.69	2803	2.14	0.884	.057	.055	.059	0.109	817.77	175	-0.024	0.005	0.021
Invarianza Stretta parziale	295	5567.632	2743	2.03	0.898	.054	.052	.056	.098	374.71	115	-0.01	0.002	0.01

Nota: India-HI senza item 11. Criteri per rigetto: $\Delta CFI (\leq .010)$, $\Delta RMSEA (\geq .015)$, $\Delta SRMR (\geq .030$ per invarianza metrica e $\geq .010$ invarianza scalare e stretta).

L'invarianza scalare parziale è stata ottenuta stabilendo vincoli di uguaglianza su 10 item nel seguente ordine: UMCS24, UMCS1, UMCS2, UMCS21, UMCS9, UMCS28, UMCS11, UMCS7, UMCS19, UMCS25. L'invarianza stretta parziale è stata ottenuta rilasciando i vincoli di uguaglianza su 10 item nel seguente ordine: UMCS25, UMCS24, UMCS28, UMCS17, UMCS3, UMCS16, UMCS9, UMCS23, UMCS13, UMCS7.

Tabella 3.10, l'invarianza scalare presenta un valore $\Delta CFI \leq .010$ e $\Delta SRMR > .010$, indici di *misfit*. Questo indica qualcosa di non invariato nel modello. È stato analizzato l'output del modello scalare, per individuare i parametri causa del *misfit*, attraverso i *modification index*. In questo modo sono stati identificati 10 intercette di items non considerabili invarianti tra i paesi, e causa dell'abbassamento di *fit* del modello. Gli items liberati dal modello sono i seguenti, in ordine: UMCS24, UMCS1, UMCS2, UMCS21, UMCS9, UMCS28, UMCS11, UMCS7, UMCS19, UMCS25. Gli items sono stati liberati uno alla volta, fino al raggiungimento di un modello di invarianza scalare con una buona *fit*. Una volta liberato questo set di items, è stato ottenuto il modello di invarianza scalare parziale, che rispetta i parametri di *fit* di Chen (2007).

L'invarianza della misura stretta prevede vincoli a livello di tutte le varianze residue tra i paesi. Questi vincoli hanno portato all'abbassamento della *fit* del modello, con valori $\Delta CFI \leq .010$ e $\Delta SRMR > .010$. Anche in questo caso, sono state identificate 10 varianze residue non invarianti tra i paesi, causa dell'abbassamento di *fit* del modello. Gli items liberati dal modello sono i seguenti, in ordine: UMCS25, UMCS24, UMCS28, UMCS17, UMCS3, UMCS16, UMCS9, UMCS23, UMCS13, UMCS7. Una volta liberato questo set di items, è stato ottenuto il modello di invarianza stretta parziale, che rispetta i parametri di *fit* di Chen (2007).

Quattro items (UMCS24, UMCS9, UMCS28, UMCS7) risultano non invarianti sia a livello scalare che a livello stretto.

3.5 Discussione

Lo studio empirico presentato rappresenta un contributo agli studi sull'equivalenza della concettualizzazione di chiamata e sulla generalizzabilità dei risultati prodotti. Come ampiamente affrontato nei precedenti capitoli, risulta fondamentale in ogni ricerca cross-culturale preoccuparsi di garantire l'equivalenza durante la verifica delle differenze cross-culturali (Hui & Triandis, 1985). Troppo spesso però i ricercatori assumono che i risultati ottenuti dalla somministrazione di uno stesso strumento a diversi gruppi coinvolti siano automaticamente comparabili (Kankaras & Moors, 2010).

Tuttavia, poiché ogni contesto culturale riflette una costellazione di molti fattori, processi e attributi, lo stesso insieme di domande o compiti può assumere significati diversi per individui appartenenti a gruppi culturali diversi, e di conseguenza misurare costrutti diversi in ogni cultura (Kankaras & Moors, 2010). In questi casi, confrontare semplicemente le medie della scala attraverso i gruppi potrebbe pregiudicare la validità delle conclusioni.

Lo scopo della ricerca presentata era quello di investigare la generalizzabilità cross-culturale della concettualizzazione di chiamata tra sette campioni appartenenti a sei nazioni (India, Turchia, Cina, Italia, Stati Uniti e Paesi Bassi), testando l'invarianza della misura della *Unified Multidimensional Calling Scale* (UMCS 28). La piena invarianza configurale e metrica della scala in questo studio dimostrano che gli item della UMCS sono ugualmente importanti nei vari paesi presi in esame per definire le sette componenti della chiamata, e che le sette componenti sono ugualmente importanti nei vari paesi per definire la chiamata.

I paesi studiati condividono quindi un'esperienza di approccio al lavoro come una chiamata, attraverso le sue componenti di passione, sacrificio, evocazione trascendente, orientamento prosociale, pervasività, significato e identità. Queste componenti riescono a sintetizzare le due prospettive del costrutto di chiamata, quella neoclassica e quella moderna, entrambe ancora presenti nelle diverse culture sia collettivistiche che individualistiche. È quindi dimostrata l'universalità della struttura e del significato del costrutto di chiamata.

Nonostante questo, la scala UMCS si conferma invariante a livello scalare e stretto in modo solo parziale. Precedenti studi sulla scala avevano già testato la scala ad un livello di invarianza della misura stretto parziale per tempo e dominio di studio (Vianello et al., 2018), e la scala si è confermata invariante a questo livello. È importante sottolineare quindi che l'utilizzo di punteggi compositi o standardizzati può portare a conclusioni fuorvianti se prima non sono stati effettuati studi di invarianza della misura. In assenza di test di equivalenza è possibile, ad esempio, interpretare un errore di misura come cambiamento di un tratto. Anche nel caso di questo studio empirico, se non fossero stati utilizzati i test che permettono di controllare bias, ci sarebbero state interpretazioni fuorvianti a livello di medie e di varianza, dato che 10 items su 28 hanno presentato bias a livello scalare o stretto.

Possibili futuri studi potranno applicare la *Unified Multidimensional Calling Scale* a più paesi o ad un campione più ampio della sola popolazione di studenti universitari di economia, per aumentare la portata di generalizzabilità dei risultati. È possibile anche immaginare la somministrazione della stessa scala allo stesso campione di paesi preso in esame in questo studio (India, Turchia, Cina, Italia, Stati Uniti e Paesi Bassi), per analizzare come altre dimensioni e fattori contestuali possano impattare la portata della

chiamata e del suo perseguimento. Nel caso della traduzione Hindi della scala, un possibile sviluppo potrebbe prevedere una nuova traduzione dell'item UMCS11, per renderne il significato equivalente a quello delle altre versioni della scala.

Bibliografia

- Alwin, D.F., & Jackson, D.J. (1981). Applications of simultaneous factor analysis to issues of factorial invariance. In D.D. Jackson and E.F. Borgotta (Eds.), *Factor analysis and measurement in sociological research: A multidimensional perspective*. Beverly Hills, CA: Sage Publications, 249-279
- Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24, 3–30.
- Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational research methods*, 1(1), 45-87.
- Bandura, A. (1982). Self-efficacy theory in human agency. *American Psychologist*, 37, 122–147.
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164-180.
- Baumeister, R. F. (1991). *Meanings of life*. New York: Guilford.
- Bellah, R. N., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. M. (1986). *Habits of the heart. Individualism and commitment in American life*. New York, NY: Harper & Row.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.

- Bentler, P. M. (1995). *EQS structural equations program manual* (Vol. 6). Encino, CA: Multivariate software.
- Bentler, P.M. (2005) EQS 6.1: Structural Equations Program Manual, Multivariate Software, Inc., Encino, CA.
- Berg, J. M., Grant, A. M., & Johnson, V. (2010). When callings are calling: Crafting work and leisure in pursuit of unanswered occupational callings. *Organization Science*, 21, 973–994.
- Blustein, D. L. (2006). *The psychology of working: A new perspective for career development, counseling, and public policy*. Mahwah, NJ: Lawrence Erlbaum.
- Bott, E. M., Duffy, R. D., Borges, N. J., Braun, T. L., Jordan, K., & Marino, J. F. (2017). Called to medicine. Physicians' experiences of a career calling. *Career Development Quarterly*, 65, 113–130.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.
- Brown, S. D., & Lent, R. W. (2016). Vocational psychology: Agency, equity, and well-being. *Annual review of psychology*, 67, 541-565.
- Bunderson, J. S., & Thompson, J. A. (2009). The call of the wild: Zookeepers, callings, and the double-edged sword of deeply meaningful work. *Administrative science quarterly*, 54(1), 32-57.
- Byrne, B., Shavelson, R., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466.

- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: a multidisciplinary journal*, 14(3), 464-504.
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural equation modeling*, 12(3), 471-492.
- Cole, D.A. and Maxwell, S.E. (1985) Multitrait-multimethod comparisons across populations: A confirmatory factor analytic approach. *Multivariate Behavioral Research*, 20, 389–417.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace
- Dik, B. J., & Duffy, R. D. (2009). Calling and vocation at work: Definitions and prospects for research and practice. *The counseling psychologist*, 37(3), 424-450.
- Dik, B. J., & Shimizu, A. B. (2019). Multiple meanings of calling: Next steps for studying an evolving construct. *Journal of Career Assessment*, 27(2), 323-336.
- Dik, B. J., Duffy, R. D., & Tix, A. P. (2012). Religion, spirituality, and a sense of calling in the workplace. In P. C Hill & B. J. Dik (Eds.), *The psychology of religion and workplace spirituality* (pp. 113–134). Charlotte, NC: Information Age.
- Dik, B. J., Eldridge, B. M., Steger, M. F., & Duffy, R. D. (2012). Development and validation of the Calling and Vocation Questionnaire (CVQ) and Brief Calling Scale (BCS). *Journal of Career Assessment*, 20, 242–263.

- Dobrow, S. R., & Tosti-Kharas, J. (2011). Calling: The development of a scale measure. *Personnel psychology*, 64(4), 1001-1049.
- Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of applied psychology*, 70(4), 662.
- Dreher, D. E., Holloway, K. A., & Schoenfelder, E. (2007). The vocation identity questionnaire: Measuring the sense of calling. *Research in the Social Scientific Study of Religion*, 18, 99–120.
- Dubin, R. (1956). Industrial worker's worlds: A study of the central life interests of industrial workers. *Social Problems*, 3, 131-142.
- Duffy, R. D., & Dik, B. J. (2013). Research on calling: What have we learned and where are we going?. *Journal of Vocational Behavior*, 83(3), 428-436.
- Duffy, R. D., Dik, B. J., Douglass, R. P., England, J. W., & Velez, B. L. (2018). Work as a calling: A theoretical model. *Journal of counseling psychology*, 65(4), 423.
- Duffy, R. D., Douglass, R. P., Autin, K. L., & Allan, B. A. (2016). Examining predictors of work volition among undergraduate students. *Journal of Career Assessment*, 24, 441– 459.
- Elangovan, A. R., Pinder, C. C., & McLean, M. (2010). Callings and organizational behavior. *Journal of Vocational Behavior*, 76, 428 – 440.
- Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93, 179-197
- Foster, E. D., & Deardorff, A. (2017). Open science framework (OSF). *Journal of the Medical Library Association: JMLA*, 105(2), 203.

- Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: An Introduction* (2nd ed.). Thousand Oaks, CA: Sage.
- Gallucci, M., Leone, L., (2017). *Modelli statistici per le scienze sociali*. Pearson.
- Gazica, M. W., & Spector, P. E. (2015). A comparison of individuals with unanswered callings to those with no calling at all. *Journal of Vocational Behavior, 91*, 1–10.
- Gerdel, S., Dalla Rosa, A., & Vianello, M. (2022). Psychometric properties and measurement invariance of a short form of the Unified Multidimensional Calling Scale (UMCS). *European Journal of Psychological Assessment*.
- Giacalone, R. A., & Jurkiewicz, C. L. (2003). Right from wrong: The influence of spirituality on perceptions of unethical business activities. *Journal of business Ethics, 85-97*.
- Greenhaus, J. H. (1971). An investigation of the role of career salience in vocational behavior. *Journal of Vocational Behavior, 1*, 209-216.
- Gustafsson, J., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28*, 407– 434.
- Hagmaier, T., & Abele, A. E. (2012). The multidimensionality of calling: Conceptualization, measurement and a bicultural perspective. *Journal of Vocational Behavior, 81(1)*, 39-51.
- Hall, D. T., & Chandler, D. E. (2005). Psychological success: When the career is a calling. *Journal of Organizational Behavior, 26*, 155–176.
- Hardy, L. (1990). *The fabric of this world: Inquiries into calling, career choice, and the design of human work*. Grand Rapids, MI: Eerdmans.

- Hirschi, A. (2011). Callings in career: A typological approach to essential and optional components. *Journal of vocational Behavior*, 79(1), 60-73.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental aging research*, 18(3), 117-144.
- Hoyle, R. H. (8). Panter, (1995). Writing about structural equation models. *Structural equation modeling: Concepts, issues, and applications*, 158-176.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, 3(4), 424.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Sage Publications, Inc.
- Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of cross-cultural psychology*, 16(2), 131-152.
- Kahn, W. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, 33, 692-724.
- Kankaras, M., & Moors, G. (2010). Researching measurement equivalence in cross-cultural studies. *Psihologija*, 43(2), 121-136.

- Kanungo, R. N. (1982). Measurement of job and work involvement. *Journal of Applied Psychology*, 67, 341-349.
- Lee, A. Y. P., Chen, I. H., & Chang, P. C. (2018). Sense of calling in the workplace: The moderating effect of supportive organizational climate in Taiwanese organizations. *Journal of Management & Organization*, 24(1), 129-144.
- Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley
- Loscocco, K. A. (1989). The interplay of personal and job characteristics in determining work commitment. *Social Science Research*, 18, 370-394.
- Marsh, H.W. (1994) Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling*, 1, 5–34.
- Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The counseling psychologist*, 33(3), 269-298.
- Maslow, A. H. (1967). A theory of metamotivation: The biological rooting of the value-life. *Journal of humanistic psychology*, 7(2), 93-127.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525–543.
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of psychological research*, 3(1), 111-130.
- Muthén, L. K., & Muthén, B. O. (2012). Mplus user's guide (1998–2012). Los Angeles, CA: Muthén & Muthén, 6.

- Newsom, J. T. (2015). Longitudinal structural equation modeling: A comprehensive introduction. Routledge.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Yarkoni, T. S. C. I. E. N. T. I. F. I. C. S. T. A. N. D. A. R. D. S. (2015). Promoting an open research culture. *Science*, 348(6242), 1422-1425.
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGrawHill.
- Peterson, C., Park, N., Hall, N., & Seligman, M. E. P. (2009). Zest and work. *Journal of Organizational Behavior*, 30, 161–172.
- Praskova, A., Creed, P. A., & Hood, M. (2015). The development and initial validation of a Career Calling scale for emerging adults. *Journal of Career Assessment*, 23, 91–106.
- Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist*, 27(4), 485-527.
- Reise, S.P., Widaman, K.F., and Pugh, R.H. (1993) Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, 114, 552–566.
- Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, 23, 51–67.
- Ruiz Quintanilla, S. A., & Wilpert, B. (1988). The meaning of working-Scientific status of a concept. *The meaning of work and technological options*, 3-14.

- Schaie, K. W., & Hertzog, C. (1985). Measurement in the psychology of adulthood and aging.
- Schuurman, D. J. (2004). Vocation: Discerning our callings in life. Grand Rapids, MI: Eerdmans.
- Serow, R. C. (1994). Called to teach: A study of highly motivated preservice teachers. *Journal of Research and Development in Education*, 27, 65-72.
- Shepard, L., Camilli, G., & Averill, M. (1981). Comparison of procedures for detecting test-item bias with both internal and external ability criteria. *Journal of Educational Statistics*, 6(4), 317-375.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, 25(2), 173-180.
- Steiger, J. H., & Lind, J. (1984). C.(1980). Statistically based tests for the number of common factors. In *the annual meeting of the Psychometric Society*.
- Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior*, 16, 282-298.
- Taris, T. W., Bok, I. A., & Meijer, Z. Y. (1998). Assessing stability and change of psychometric properties of multi-item concepts across different situations: A general approach. *The Journal of Psychology*, 132(3), 301-316.
- Thompson, J. A., & Bunderson, J. S. (2019). Research on work as a calling... and how to make it matter. *Annual Review of Organizational Psychology and Organizational Behavior*, 6, 421-443.

- Uchida, Y., Norasakkunkit, V., & Kitayama, S. (2004). Cultural constructions of happiness: theory and empirical evidence. *Journal of happiness studies*, 5(3), 223-239.
- Van de Vijver, F. J. (1998). Towards a theory of bias and equivalence. 50967, 3, 41-65.
- Van de Vijver, F., & Leung, K. (1997). Methods and data analysis of comparative research. *Handbook of cross-cultural psychology*, 1, 257-300.
- Van de Vijver, F., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment. *European review of applied psychology*, 47(4), 263-279.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.
- Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, 13(4), 411-417.
- Vianello, M., Dalla Rosa, A., & Gerdel, S. (2022). Career calling and task performance: The moderating role of job demand. *Journal of Career Assessment*, 30(2), 238-257.
- Vianello, M., Dalla Rosa, A., Anselmi, P., & Galliani, E. M. (2018). Validity and measurement invariance of the Unified Multidimensional Calling Scale (UMCS): A three-wave survey study. *Plos one*, 13(12), e0209348.
- Vianello, M., Galliani, E. M., Rosa, A. D., & Anselmi, P. (2020). The developmental trajectories of calling: Predictors and outcomes. *Journal of Career Assessment*, 28(1), 128-146.

- Wang, J., & Wang, X. (2012). Structural equation modeling with Mplus: methods and applications. Wiley.
- Weber, M. (1930). The Protestant ethic and the spirit of capitalism (T. Parsons, Trans.). New York: Scribner. (Original work published 1904/05)
- Weston, R., & Gore Jr, P. A. (2006). A brief guide to structural equation modeling. *The counseling psychologist*, 34(5), 719-751.
- Wrzesniewski, A. (2012). Callings. In K. S. Cameron & G. Spreitzer (Eds.), Handbook of positive organizational scholarship (pp. 45–55). New York, NY: Oxford University Press.
- Wrzesniewski, A., & Dutton, J. E. (2001). Crafting a job: Revisioning employees as active crafters of their work. *Academy of management review*, 26(2), 179-201.
- Wrzesniewski, A., McCauley, C., Rozin, P., & Schwartz, B. (1997). Jobs, careers, and callings: People's relations to their work. *Journal of research in personality*, 31(1), 21-33.