



**Assessment e Serious Game:
Una nuova strada per valutare le Soft Skill?**
Analisi critica di uno progetto pilota

Dott.ssa Barbara Benincasa

**DIPARTIMENTO DI STUDI UMANISTICI
Università di Napoli Federico II**

**Dottorato di Ricerca in
HUMAN MIND AND GENDER STUDIES - XXIX Ciclo**

Supervisore: Prof. Davide Marocco

Introduzione

Il presente studio, di natura multidisciplinare, si muove nell'ambito della valutazione delle soft skill, o competenze personali, realizzata con strumenti innovativi e tecnologici. L'obiettivo che si pone è duplice:

1. Presentare uno strumento, tecnologicamente e metodologicamente innovativo, per l'assessment psicologico, in particolare per la valutazione delle competenze di gestione dei conflitti.
2. Indentificare tecniche psicometriche valide per lo studio della bontà di questo nuovo strumento di valutazione, che esce dai confini dell'assessment tradizionale.

La ricerca propone un innovativo strumento di valutazione delle competenze di negoziazione, il serious game ENACT, gioco nato nell'ambito di un progetto finanziato dalla Comunità Europea, come prodotto alternativo ai tradizionali metodi "Carta e Matita". Con questo nuovo strumento, si è indagata la possibilità di integrare le teorie e le metodologie della valutazione psicologica con i Serious games, ossia giochi nati e utilizzati per scopi educativi e formativi. L'elaborato descrive la realizzazione di ENACT, partendo dai costrutti teorici di Rahim sulla gestione dei conflitti e dal test da lui realizzato, il ROCI II, in quanto punto di partenza dell'intero lavoro svolto. Inoltre, descrive le caratteristiche di questo nuovo test e i risultati ottenuti nelle diverse fasi della sua implementazione, finalizzati alla verifica della sua validità e attendibilità. Lo scopo finale è verificare quanto il serious game può rappresentare una forma parallela del test d'origine nella valutazione degli stili di negoziazione. *Questo specifica parte del lavoro è stato svolta con il gruppo del laboratorio NAC, Natural and Artificial Cognition Lab (www.nac.unina.it), dell'università di Napoli Federico II, nel biennio 2014-2015, sotto il coordinamento dell'Università di Plymouth e la collaborazione di altri 4 partner europei.*

Per il soddisfacimento del secondo obiettivo sono stati utilizzati i dati di questo studio pilota per un'analisi critica dei risultati finalizzata a verificare la bontà della nuova tecnica, e in particolare per identificare un metodo psicometrico valido per il confronto dei due strumenti, che sono di fatto radicalmente diversi.

L'analisi effettuata alla fine del progetto, infatti, se da una parte ha mostrato alcuni elementi positivi, dall'altra ha fatto sorgere molte domande di natura statistica, legate soprattutto ai metodi utilizzati per l'analisi dei dati, vista la diversa natura dei due test, carta e matita vs serious game, e le differenze dei dati, sia in termini di numerosità che di tipologia. In altre parole, l'analisi critica effettuata utilizza tecniche della teoria classica dei test per trovare un metodo valido i) per rappresentare la ricchezza e la varietà di informazione offerta da ENACT, ii) per confrontare i due strumenti su un piano squisitamente psicometrico. Il risultato finale è rappresentato da una serie di considerazioni e suggerimenti volti a migliorare questo nuovo test di valutazione, che sebbene sia ancora in fase prototipale e presenta diverse difficoltà, rappresenta comunque un valido e innovativo strumento di assessment soprattutto in riferimento ai cambiamenti sociali e culturali a cui si sta assistendo negli ultimi decenni. *Questo seconda parte del lavoro rappresenta il contributo personale realizzato nell'ultimo anno di dottorato.*

L'elaborato, che descrive l'intero lavoro svolto, si articola in 4 capitoli principali. Il **primo** fornisce un'introduzione ai concetti di assessment, soft skill e serious game, i tre tempi di ricerca sui cui è incentrato il progetto, descrivendo l'interazione tra questi, con un riferimento particolare all'ultima evoluzione dell'assessment psicologico spinto ormai sempre più verso tecnologie più sofisticate e innovative. Nel **secondo** capitolo sarà descritta la teoria di Rahim e il test da lui sviluppato sulla gestione dei conflitti, il ROCI II, in quanto punto di partenza dell'intero lavoro. Il **terzo** capitolo rappresenta la parte centrale di questo studio in quanto descrive le diverse fasi della ricerca, ed è articolato in due sezioni distinte, una per ogni obiettivo preposto. In particolare una prima sezione si focalizza sul progetto ENACT con specifico riferimento agli assunti teorici, alle modalità di realizzazione dello strumento e alle analisi dei dati, finalizzati a dimostrare il parallelismo delle due tecniche. Molto importante in questa sezione, è l'analisi critica dei risultati del progetto, nella quale sono individuate tutte quelle differenze che impattano sui risultati ottenuti e che hanno permesso l'identificazione di una serie di interrogativi sulla validità statistica delle analisi, interrogativi che hanno poi guidato la seconda parte della ricerca. In questa seconda parte, pertanto, utilizzando i dati raccolti nel corso del progetto, si introduce l'analisi esplorativa condotta al fine di verificare se le

tecniche utilizzate nell'ambito del progetto siano le più adatte allo scopo e soprattutto si pone l'obiettivo più alto di identificare e descrivere un metodo psicometrico utile al confronto dei due test che, come abbiamo detto, seppur nati dallo stesso approccio teorico, sono in realtà completamente differenti. L'elaborato si conclude con **l'ultimo** capitolo dedicato alla discussione dei risultati ottenuti e ai suggerimenti per i successivi studi.

In conclusione, ENACT, strumento pioneristico nato nell'ambito di un progetto finanziato dalla Comunità Europea e sviluppato nel corso di due anni, viene qui considerato come un primo passo verso la realizzazione di test evoluti, basati cioè sull'utilizzo di serious game come forma di valutazione psicologia, e quindi più orientati a rispecchiare i cambiamenti sociali e culturali dei nostri giorni. Sebbene i suoi risultati sono promettenti, la ricerca mostra che il viaggio è ancora lungo....

Sommario

INTRODUZIONE	1
1 OVERVIEW	5
1.1 L'ASSESSMENT PSICOLOGICO	6
1.1.1 L'EVOLUZIONE DEI TEST PSICOLOGICI	9
1.1.2 CARATTERISTICHE E FUNZIONI DEI TEST	11
1.1.3 APPLICAZIONE E CLASSIFICAZIONE DEI TEST	15
1.2 LE SOFT SKILL	19
1.2.1 COMPETENZA VS COMPETENZA TRASVERSALE	22
1.2.2 I MODELLI DELLE COMPETENZE	27
1.2.3 LE SOFT SKILL E LE NUOVE TECNOLOGIE	31
1.3 I SERIOUS GAME	33
1.3.1 I GIOCHI DI RUOLO PER LA FORMAZIONE DELLE SOFT SKILL	37
1.3.2 I GIOCHI DI RUOLO VS SERIOUS GAME	40
1.3.3 I SERIOUS GAME COME STRUMENTI DI ASSESSMENT	44
2 IL PUNTO DI PARTENZA: IL ROCI E LA TEORIA DI RAHIM	47
2.1 RAHIM E LA GESTIONE DEI CONFLITTI	48
2.2 IL ROCI II	52
2.3 PERCHÉ IL MODELLO DI RAHIM	53
3 LA RICERCA	56
3.1 OBIETTIVI	56
3.2 ENACT: UN NUOVO STRUMENTO DI VALUTAZIONE PSICOLOGICA	58
3.2.1 OBIETTIVI DEL PROGETTO	58
3.2.2 ENACT ASSESSMENT: LO STRUMENTO	60
3.2.3 LE CARATTERISTICHE PRINCIPALI DI ENACT	62
3.2.4 LA VALIDAZIONE	64
3.2.5 L'IMPLEMENTAZIONE	66
3.2.6 ENACT – ROCI II: DISCUSSIONE CRITICA DEI RISULTATI	70
3.3 ANALISI ESPLORATIVA SUI RISULTATI DEL PROGETTO	76
3.3.1 METODOLOGIA D'ANALISI	77
3.3.2 PREPARAZIONE DEI DATA SET	83
3.3.3 ANALISI 1	86
3.3.4 ANALISI 2	90
3.3.5 ANALISI 3	92
3.3.6 ANALISI 4	98
4 CONCLUSIONI	101
ALLEGATO 1 - ROCI II FORM C	104
ALLEGATO 2 - ENACT ASSESSMENT – SCREENSHOT SCENARI	105
BIBLIOGRAFIA	112

1 Overview

Negli ultimi decenni si stanno verificando una serie di fenomeni di carattere culturale, sociologico, tecnologico ed economico che stanno impattando in maniera diretta e visibile su tutti gli aspetti di vita degli individui, sia personali che professionali. La crescita individuale, il modo di lavorare, le modalità di relazionarsi, ad esempio, sono state totalmente trasformate sia dall'avvento di internet e delle nuove tecnologie, che da un nuovo modo di definire e valutare l'essere umano e le sue capacità. Parole come e-learning, tecnologie innovative, social network, soft skill, testing, fanno ormai parte del linguaggio comune, non senza però confusione e contraddizioni. Questo perché è evidente, in una modalità del tutto peculiare di questo momento storico, l'interesse sempre più esteso ad indagare le capacità specifiche degli individui, di comprendere come potenziarle, attraverso l'utilizzo di strumenti di formazione tecnologicamente sempre più sofisticati e innovati, al fine di consentire una maggiore inclusione sociale e professionale, che abbracci e rappresenti i cambiamenti radicali, specifici di quest'epoca. Questo fenomeno, che caratterizza tutto il mondo industrializzato, in Europa è ben rappresentato dagli sforzi della Comunità Europea che, già nei programmi di finanziamento precedenti, ma soprattutto con Horizon 2020¹, ha chiaramente definito quali devono essere le caratteristiche dei cittadini della nuova comunità, determinando di fatto una crescita ingente e costante di impegno da parte di organizzazioni, enti e singoli nella creazione di un sistema unico che tenga conto di tutte queste dimensioni e che, attraverso uno sforzo multidisciplinare, consenta di raggiungere gli obiettivi prefissati, superando gli ostacoli e le difficoltà tipiche del cambiamento.

L'elemento chiave di questa strategia² è l'individuo, che viene delineato come persona consapevole delle proprie capacità e competenze, attento e coinvolto

¹ Horizon 2020 è lo strumento di finanziamento alla ricerca scientifica e all'innovazione della Commissione europea che ha un budget stanziato tra i più alti del mondo: quasi 80 miliardi di euro, per 7 anni (2014 al 2020).

²*La strategia Europa 2020 punta a rilanciare l'economia dell'UE nel prossimo decennio. In un mondo che cambia l'UE si propone di diventare un'economia intelligente, sostenibile e solidale. ...[per] conseguire elevati livelli di occupazione, produttività e coesione sociale. In pratica, l'Unione si è posta cinque ambiziosi obiettivi – in materia di occupazione, innovazione, istruzione, integrazione sociale e clima/energia – da raggiungere entro il 2020.* Tratto dal sito web http://ec.europa.eu/europe2020/index_it.htm

nell'ambiente e nella società in cui vive, preparato e pronto nel proprio lavoro, così da rendere l'Europa un continente competitivo, innovativo e sicuro. Per rendere questo non solo un'ideale, ma qualcosa di fattibile e quantificabile, negli ultimi anni un numero enorme di ricerche, studi e applicazioni, in una serie di ambiti diversificati, hanno posto proprio l'individuo al centro della loro indagine. Prima fra tutte è emersa la necessità di capire la persona e le proprie peculiarità, al fine di valutarle e meglio formarle in vista di questo nuovo traguardo. La definizione, la formazione e la valutazione delle competenze individuali, e in particolare di tutte quelle competenze trasversali, non legate cioè ad un sapere appreso, ma connesse alla capacità dell'uomo di muoversi attivamente e funzionalmente nel mondo, le cosiddette soft skill, sono pertanto diventate punto di interesse principale, insieme a tutti quegli sviluppi della moderna tecnologia che rendono questi processi più semplici, immediati e immersi nella società attuale. Nasce così un nuovo campo di indagine, multidisciplinare, vasto e complesso, che affronta questi temi, proponendo nuovi strumenti e tecniche: il campo della valutazione e formazione delle competenze individuali mediante strumenti tecnologici avanzati. È proprio in questo settore che si muove la presente ricerca, con l'obiettivo di presentare uno degli ultimi prodotti realizzati, analizzandone le criticità e gli aspetti positivi e individuando quegli elementi di successo che fanno di strumenti come questi, tecniche essenziali per il prossimo futuro.

Ma prima di tutto, è necessario delineare quest'ambito di studio e definire i principali concetti su cui si fonda. Questo capitolo si focalizza pertanto su tre dei temi centrali di questo ampio e articolato settore d'indagine: assessment, soft skill e serious game, e ne descrive in particolare l'interazione per chiarire quella che può essere considerata l'ultima evoluzione dell'assessment, ossia la valutazione come forma di gioco digitale.

1.1 L'Assessment psicologico

Il termine **assessment** deriva dal latino *assidere* "sedere come giudice" ed è un termine inglese che significa "valutare, stimare, giudicare". In ambito psicologico, *si riferisce ad una valutazione globale e differenziale del paziente, nell'unicità e complessità psicologica che lo caratterizza, considerando anche le sue risorse e i*

suoi limiti. Si tratta pertanto di una raccolta di informazioni particolarmente ampia volta a conoscere la storia e le caratteristiche del problema presentato dalla persona, le ricadute sulla sua vita quotidiana, e la sua storia personale.

Nell'ampia letteratura sul tema è possibile ritrovare tuttavia diverse definizioni del termine *assessment*, a seconda degli ambiti di applicazione e delle prospettive di riferimento: nelle psicoterapie ad orientamento dinamico, ad esempio, l'*assessment* è visto come quel processo utile a determinare la presa in carico di un soggetto, o la possibilità che esso benefici del trattamento, mentre negli approcci cognitivi comportamentali, viene utilizzato nelle fasi iniziali per identificare gli elementi che potrebbero influenzare l'esito di una terapia. Il *clinica assessment* (o *valutazione clinica*), si riferisce invece ad una misurazione e valutazione sistematica di fattori psicologici, sociali e biologici di una persona che presenta un probabile disturbo mentale (Barlow & Durand, 1995). Nella psicologia del lavoro, infine, è utilizzato sia per la selezione del personale che per la valutazione delle competenze dei lavoratori, al fine di creare percorsi di formazione personalizzati.

Indipendentemente dalla modalità d'uso, l'*assessment* rientra in un ambito di studio più ampio, la **psicodiagnosi**, definita come quella disciplina che si occupa della *valutazione attraverso un processo conoscitivo complesso ed integrato che ha tra le sue finalità il prendere una decisione che prelude ad un trattamento specifico* (Scafidi Fonti, 2000; Parker, 1999). Questo processo è caratterizzato dall'uso integrato di molteplici strumenti quali questionari, inventari di personalità, batterie e tecniche testistiche (psicometriche e proiettive), colloqui clinici, esami neuropsicologici e valutazioni osservative. Naturalmente, il tipo di tecniche e strumenti variano, in base al contesto e allo scopo della valutazione, all'età dei soggetti ed al tipo di difficoltà che presentano, oltre che all'orientamento teorico e alla formazione specialistica del valutatore.

È importante sottolineare che il processo diagnostico, così definito, si differenzia in funzione della sua finalità e degli ambiti di applicazione. La prima distinzione importante da fare è tra *valutazione psicologica* e *psicopatologica*.

La *valutazione "classica"* è solitamente di tipo nosologico e psicopatologico, ed è effettuata attraverso un colloquio clinico ed anamnestico, eventualmente integrato

dalla somministrazione di scale di rilevazione della sintomatologia psichiatrica. Espressione specifica dell'ambito clinico-psicopatologico, l'attività di diagnostica è, in questo contesto, uno strumento essenziale per definire una diagnosi psichiatrica, spesso secondo i criteri nosografici delle classificazioni internazionali del DSM-IV o dell'ICD-10. La diagnostica psicologica invece può essere di tipo più ampio: oltre che alla rilevazione di sintomatologia psicopatologica, infatti, può essere riferita anche alla valutazione di aspetti e processi della personalità, alla valutazione di atteggiamenti, modalità relazionali, livello e tipologia di competenze cognitive, struttura di personalità, ecc. Diversamente dalla diagnosi "classica", quella psicologica quindi non si limita al riconoscimento ed alla classificazione dei sintomi o all'inquadramento di una malattia ma, prende in carico l'individuo nella sua complessità e unicità, ed esamina e studia le sue risorse, le modalità individuali di affrontare una crisi o un problema, si propone in altre parole di giungere ad una comprensione psicologica che passa necessariamente attraverso la condivisione emotiva e cognitiva di aspetti profondi del sé. L'utilizzo dei test psicologici completa ed integra le informazioni raccolte durante i colloqui e permette di valutare le funzioni o le caratteristiche specifiche di personalità, consentendo di avere in tempi relativamente brevi, una conferma o disconferma delle osservazioni cliniche.

L'uso dei test, e più in generale dei presidi psicodiagnostici, rappresenta quindi una delle modalità di valutazione più versatili del lavoro psicologico. Infatti, oltre ad essere utilizzati come momento di avvio in una relazione terapeutica, i test possono avere un ruolo funzionale ben articolato anche in molte altre fasi, sia come fotografie in itinere, che come follow-up dopo la conclusione della relazione con il paziente. Per la loro stessa natura, inoltre, i test vengono utilizzati anche per finalità estranee all'ambito clinico, ma maggiormente centrate sulla misurazione e valutazione di specifici aspetti, all'interno di ambiti più mirati e finalizzati allo scopo, come ad es. l'orientamento, la formazione, l'istruzione, la sanità, le organizzazioni, la pubblicità, il marketing. In tutti questi ambiti applicativi, possono essere utilizzati un numero cospicuo di test, che si differenziano per tipologia e finalità, caratteristiche che si sono evolute e affinate soprattutto nel corso degli ultimi due secoli.

1.1.1 L'evoluzione dei test psicologici

Sebbene l'interesse per la misurazione delle caratteristiche psichiche dell'individuo si può far risalire a qualche secolo prima di Cristo (Chuan-Thing, 1980; Boncori, 1993), la nascita dei test, nella loro accezione più moderna, è della seconda metà del diciannovesimo secolo, quando Galton, influenzato dall'opera di Darwin sulle Origini della Specie, iniziò uno studio sugli aspetti ereditari dei superdotati, costruendo i primi reattivi per misurare le capacità cognitive. Fondò nel 1882 il primo centro dei test a Londra, dove studiò la relazione tra intelligenza e capacità sensoriali su circa 17.000 persone, convinto che la discriminazione sensoriale potesse essere una misura delle capacità intellettive di un individuo. Sulla stessa scia di pensiero nacquero i primi lavori di Cattell (1860-1944) che come Galton sosteneva l'idea di un legame tra l'intelligenza e capacità sensoriale. Tuttavia, Cattell si concentrò solo sulle funzioni semplici, convinto dell'impossibilità di misurare le funzioni umane complesse, con il risultato di creare di fatto solo reattivi psicofisici, che valutavano aspetti secondari come la velocità di movimento o l'acutezza della vista. Questa scelta fece nascere molte riserve e perplessità, sia da un punto di vista teorico che metodologico, tanto da spingere alcuni psichiatri europei, come Kraepelin e Jung, ad allontanarsi da questo approccio per costruire ed utilizzare dei test specifici per l'ambito psichiatrico.

È nel 1905 che compare il primo test di intelligenza ad opera di Binet, che contrariamente all'americano Cattell, sosteneva l'idea di poter misurare le funzioni complesse grazie alla notevole differenza esistente tra gli individui. Costruì così la prima Scala di Intelligenza, con reattivi specifici per la memoria, l'attenzione, la comprensione, l'immaginazione, che venne successivamente modificata e ampliata nella versione 2 del 1908 e nella versione 3 del 1911. L'elemento centrale e innovativo della teoria di Binet è rappresentato dal suo concetto di età mentale, ossia l'età rilevabile dal numero di prove del suo test superate in maniera positiva, e rapportata all'età media dei bambini della stessa età. I lavori che presero spunto da questi studi portarono all'introduzione del quoziente di intelligenza (QI), proposto per la prima volta da Terman nel 1916. Il QI, rappresentato dal rapporto tra età cronologica (EC) ed età mentale (EM) moltiplicato per 100, comparve a seguito della pubblicazione da parte di Terman della scala Stanford-Binet (così denominata perché

fu sistemata alla Stanford University). La definizione del concetto di QI rappresentò una spinta enorme in questo campo, che se da una parte permise di ampliare e approfondire gli studi sull'intelligenza, dall'altro determinarono una serie di modifiche agli aspetti centrali della teoria di Terman. Innanzitutto, il concetto di età mentale fu sostituito da *età dello sviluppo* (Gesell, 1925), così come gli stessi studi sull'intelligenza passarono lentamente a considerarla non più come qualcosa di unico e totale, ma come l'insieme di vari tipi di intelligenze (vedi paragrafo 1.2.1), e di conseguenza anche i test utilizzati per misurarla subirono profondi mutamenti, passando da reattivi basati esclusivamente sulle capacità verbali a reattivi non verbali (come ad es. le Matrici proiettive di Raven, del 1938). Ma il vero passaggio all'uso dei test in ambiti diversi da quello della ricerca fu determinato da un duplice fenomeno: da una parte il moltiplicarsi, in tutto il mondo, di studi e di strumenti per la misurazione dell'intelligenza e dall'altro le necessità specifiche legate all'inizio del primo conflitto mondiale che portò l'esercito americano ad adottare, per la prima volta, un approccio psicometrico per la selezione delle reclute. A tal fine, un gruppo di psicologi sviluppò due scale di misura³ del livello intellettuale (unico elemento all'epoca misurabile), che per motivi organizzativi e dimensionali, furono pensate specificatamente per la somministrazione di gruppo, altra importante innovazione. L'utilizzo da parte dell'esercito americano, non solo aprì la strada alla creazione di numerosi reattivi applicabili in diversi settori, ma ancor di più diede l'avvio all'utilizzo dei test nelle scuole, tutt'oggi ancora molti utilizzati negli USA, sebbene non senza critiche e contestazioni. In generale, il loro utilizzo nei vari ambiti, soprattutto in America, fu cospicuo fino alla seconda metà del XX secolo. A partire dal 1960, tuttavia, ci fu un brusco rallentamento dovuto: i) all'interesse, in ambito psicologico, a temi maggiormente legati alle specificità dell'individuo nei vari settori di vita e di lavoro; ii) al cambiamento della figura dello psicologo, che cominciò ad assumere un ruolo sempre più terapeutico, con l'esplosione, dopo la seconda guerra mondiale, dell'interesse verso la psicologia clinica, che di fatto eliminava dalla sua pratica l'utilizzo di test psicodiagnostici, ad esclusione dei reattivi proiettivi (vedi paragrafo 1.1.3); e iii) alla crisi dell'industria dei test, che dopo l'iniziale entusiasmo

³ Army alpha test e Army beta test: il primo per persone che sapevano leggere e scrivere, il secondo per gli analfabeti e quindi basato su item non verbali.

generato dall'idea di poter misurare l'intelligenza umana, portò ad un uso e abuso di queste tecniche spesso realizzate senza sostegno teorico e metodologico, ma solo per motivi esclusivamente pecuniari, che ben presto dimostrarono la loro scarsa robustezza. L'insieme di queste problematiche e la diatriba sempre più aspra tra psicologi clinici e psicometrici portò nel corso degli anni all'individuazione di numerose criticità con il risultato di una svalutazione generale di entrambi gli approcci. Tuttavia, dopo un momento di stallo, si è osservata negli ultimi decenni una graduale ripresa dell'interesse in questo campo, dovuta in particolare all'immediata utilità che questi strumenti offrono, soprattutto in contesti che necessitano di risultati immediati in poco tempo, come ad esempio le organizzazioni. In ogni caso, qualsiasi sia l'ambito di applicazione quello che l'evoluzione storica ha determinato è che, per avere risultati affidabili, è fondamentale utilizzare solo dei **buoni test**, costruiti cioè secondo dei rigidi criteri teorici e metodologici, in maniera adeguata alla funzione e al contesto di riferimento.

1.1.2 Caratteristiche e funzioni dei test

Oggi il **test** è definito come uno *strumento di misura costruito sulla base di una serie di elementi teorici e metodologici che ne rappresentano la robustezza e l'affidabilità*. Gli elementi fondamentali di un buon test sono diversi. Innanzitutto, gli **item** che possono essere, a seconda del tipo di test, domande, prove grafiche o motorie, o prestazioni di vario genere. Elemento centrale di ogni reattivo, gli item devono necessariamente essere rappresentativi di una singola variabile o costrutto psicologico, oltre ad essere uguali per tutti i soggetti ed inalterabili. Di estrema importanza infatti è anche la **modalità di somministrazione**, che deve rimanere identica e seguire le stesse regole, indipendentemente dal campione oggetto di studio. In altre parole, un test deve poter essere utilizzato usando sempre lo stesso setting, le stesse istruzioni, gli stessi item, con tutti i tipi di soggetti. La **raccolta e codifica dei dati**, inoltre, deve essere obiettiva e ripetibile, ossia scevra da qualsiasi influenza esterna, come ad esempio le teorie dell'esaminatore, e quindi riproducibile in ogni contesto, così da fornire garanzie sulla omogeneità del test. Attualmente, quest'aspetto del testing è facilitato da numerosi software sviluppati specificatamente per questo scopo. La somministrazione e la codifica dei dati sono alla base della **standardizzazione**, altro aspetto centrale nella costruzione di un test. Si tratta di un

procedimento matematico che permette di riportare tutti i punteggi su una curva di distribuzione normale, notoriamente detta Curva di Gauss, caratterizzata da una distribuzione con media zero e varianza pari a 1, finalizzata alla generalizzazione dei risultati attraverso il confronto dei punteggi ottenuti da un campione con quelli di una popolazione rappresentativa, a cui il campione esaminato appartiene. La standardizzazione di un test, garantita da una buona *taratura*, che si effettua nella fase di costruzione, e da un rigoroso metodo di somministrazione, è obbligatoriamente riportata nei singoli manuali di ciascun reattivo, insieme ai valori di attendibilità e validità, considerate di fatto le due misure principali per identificare la bontà di uno strumento. L'**attendibilità** è definita come *la capacità di un test di misurare una dimensione psicologica indipendentemente dell'errore casuale di misura*. Questo significa che un test è attendibile se non è soggetto alla variabilità di fattori esterni, quali le variazioni momentanee dei soggetti, o il contesto di applicazione, o elementi casuali di varia natura. Esistono molti modi per misurare l'attendibilità, legati soprattutto alle modalità di somministrazione. Nei test ripetuti, ad esempio, si utilizzano le tecniche del test-retest o dei test paralleli, mentre per test con un'unica somministrazione, si ricorre al metodo split-half o al metodo della consistenza interna. Nelle scienze sociali, la formula più comune per misurare l'attendibilità è l'Alpha di Crombach, che si basa sulla coerenza interna degli item, ossia il grado di accuratezza degli item nel misurare tutti la stessa caratteristica. La **validità**, invece è definita come la capacità di un test di misurare quello che doveva misurare. Si tratta di *“un giudizio complessivo della misura, in cui prove empiriche e principi teorici supportano l'adeguatezza e l'appropriatezza delle conclusioni basate sui punteggi al test”* (Messick, 1989). Anche in questo caso, esistono diversi tipi di validità, ognuno caratterizzato da un diverso modo di approcciare al tema. Nella tassonomia tradizionale, si distingue tra validità di facciata e validità relativa al contenuto (vedi Figura 1), che a sua volta si articola in diverse modalità di misura.

1. La **Validità di Facciata** indica il grado in cui il test sembra essere una misura sensata del costrutto, ed è basata sulla possibilità di comprendere il costrutto misurato dalla lettura degli item utilizzati. L'aspetto positivo della validità di facciata è che agisce sensibilmente sulla motivazione dei soggetti a completare il

test, ma ne determina anche la possibilità di modificare le risposte, e questo è sicuramente l'aspetto più criticato.

2. La **Validità di Contenuto** indica quanto il campione di item utilizzato rappresenta il costrutto oggetto di analisi. In altre parole, gli item del test devono coprire l'intera definizione teorica del costrutto in esame, e riuscire a cogliere gli aspetti più importanti per la sua misurazione. Essendo molto difficile da valutare, non esistono metodi quantitativi, ma solo qualitativi come ad es. il parere di esperti.
3. La **Validità di criterio** è il grado di corrispondenza tra la misura osservata e una variabile esterna, diversa dal costrutto originario, che si assume come criterio di riferimento per un confronto. La logica sottostante è un criterio esterno possa essere considerato una buona approssimazione della variabile di interesse. Essa può essere di due tipi: predittiva o concorrente, a seconda che il punteggio del test risulti utile nello predire o stimare il punteggio della variabile criterio. La **validità predittiva** si misura rapportando il punteggio ottenuto ad uno specifico test con un criterio esterno, successivo ed indipendente alla misurazione (ad esempio, la valutazione a posteriori del rendimento lavorativo degli individui selezionati). La rilevazione del criterio è posteriore al test, e prevede un intervallo temporale tra la somministrazione e la rilevazione del criterio. Per la validità **concorrente** invece il punteggio ottenuto al test è rapportato ad un criterio rilevato nello stesso momento (ad esempio, l'osservazione dei candidati in una situazione di interazione). Entrambe le misure sono basate sulla correlazione tra i punteggi ottenuti al test e quelli del criterio esterno.
4. La **Validità di costrutto** indica il grado in cui uno strumento misura il costrutto che si intende misurare. Poiché i costrutti sono astrazioni concettuali, questo tipo di validità bisogna derivarla in maniera indiretta, e quindi essa può essere **convergente**, nel caso il test concordi con altre misure della stessa variabile, o **divergente**, quando invece discosta o più tecnicamente, non correla, con strumenti che misurano costrutti diversi. Le tecniche più utilizzati per questa misura sono: i gruppi contrapposti, le Matrici Multi-Tratto-Multi-Metodo (MTMM), le Matrici Multi-Tratto-Multi-Item (MTMI), l'Analisi Fattoriale.



Figura 1: Tipi di validità di un test

Ogni test quindi deve essere costruito in maniera rigorosa e soprattutto deve necessariamente essere valido e attendibile per uscire dall'ambito della ricerca ed essere applicato nei diversi contesti d'uso. Un altro aspetto fondamentale, in questo senso, è la scelta del test giusto per l'obiettivo preposto, in altre parole la scelta del test in base alla sua funzionalità. È importante sottolineare che a seconda della funzione, i test devono soddisfare uno o più criteri di validità. Un test può avere tre diverse funzioni:

1. **Funzione predittiva:** ossia la capacità di predire in che modo i soggetti si comporteranno in situazioni diverse da quella d'esame. Più un test è buono, maggiore sarà la capacità di stimare la performance dei soggetti per quella specifica caratteristica al di fuori del setting artificiale, in contesti di vita quotidiana. Un esempio è dato dai test per la selezione del personale, che hanno l'obiettivo di prevedere la riuscita lavorativa dei soggetti una volta assunti. Appare evidente che questi tipi di test devono possedere una buona validità predittiva basata su una identificazione chiara del criterio di riferimento, che nel caso esemplificato potrebbe essere rappresentato dal numero di soggetti che hanno mantenuto il lavoro dopo il periodo di prova.
2. **Funzione valutativa:** ossia la capacità di misurare in maniera diretta il rendimento di un soggetto in una particolare situazione. Un esempio classico di test di questo genere sono i test d'esame all'università. L'aspetto importante è la loro validità di contenuto, ossia la capacità di coprire in maniera rilevante l'intero ambito di nozioni che si vuole valutare. Sebbene sembri molto semplice da realizzare, in realtà è un processo molto complesso perché influenzato da

elementi esterni, quali ad esempio, nel caso specifico, la competenza e la conoscenza dell'esaminatore che ha preparato il test, o le differenze tra i gruppi di studenti o le caratteristiche dell'università stessa. In altre parole, in questi casi, non sempre quel che si valuta è identico in ogni contesto, e spesso questi tipi di test sono strettamente legati all'ambito per cui sono stati costruiti.

3. **Funzione diagnostica:** ossia la capacità di discriminare e descrivere le caratteristiche psicologiche di un individuo, delineandone le caratteristiche, le anomalie, gli elementi fondamentali, al fine di identificare ciò che può essere considerato fuori dalla "norma". I test diagnostici sono utilizzati nella clinica, in contesti legali e sanitari, oltre che nella ricerca, ma in ogni caso, l'elemento essenziale è la bontà degli strumenti utilizzati che non si limita alla validità predittiva o di contenuto, come nei casi precedenti, ma che deve soddisfare tutti (o quasi) i criteri sopra descritti.

1.1.3 Applicazione e classificazione dei test

Come abbiamo visto, la costruzione dei test è strettamente connessa ad una serie di aspetti fondamentali che ne determinano la bontà e la loro applicazione. La relativa facilità d'uso e il riscontro immediato dei risultati ne hanno determinato un interesse notevole in tutti i settori, aumentando pertanto in breve tempo gli ambiti di applicazione. La ricerca clinica è sicuramente quello di maggiore rilevanza teorica, intendendo con questo termine sia gli studi di validazione e taratura degli strumenti, sia lo studio e l'analisi dei processi psicologici e psicopatologici mediante reattivi psicodiagnostici. In entrambi questi casi, solitamente, l'intento è quello di ottenere inferenze generalizzabili attraverso la raccolta di dati aggregati di campioni sufficientemente ampi e rappresentativi, come ad esempio, nella ricerca sui fattori di rischio di determinate problematiche. Anche in ambito educativo si utilizzano test, sia per la valutazione del livello intellettuale, delle capacità cognitive e delle competenze metacognitive, sia per l'identificazione di eventuali problematiche emotive e cognitive correlate a situazioni di problematicità o disabilità psicofisica, sia per la valutazione della performance o dell'apprendimento. Un altro ambito di forte utilizzo è la valutazione psicologica a scopo legale, peritale o assicurativo, soprattutto per la realizzazione di perizie civilistiche o penalistiche come i CTU o

CTP, valutazioni di idoneità per le adozioni e gli affidamenti, valutazioni di idoneità alla guida, etc.. Sebbene i test e le modalità di somministrazione siano le stesse degli altri settori di applicazione, questo tipo di valutazione è caratterizzata da una diversa prospettiva di indagine dove il soggetto valutato non è più il "referente finale" della valutazione stessa, la quale invece è solitamente richiesta da un ente esterno, come un' autorità giudiziaria o una commissione medico-legale. Questo diversa inquadratura modifica di fatto alcuni degli assetti sia formali che deontologici della procedura psicodiagnostica, che per la sua complessità e delicatezza, anche sotto un profilo normativo e professionale, implica la necessità di un'adeguata preparazione e certificazione per il suo svolgimento. Si parla invece di valutazione (assessment) neuropsicologica quando la diagnosi è finalizzata all'individuazione di eventuali difficoltà neurocognitive, derivanti da deficit e lesioni encefaliche (post-traumatiche, post-ictus, legate a tumori, etc.) o da forme di demenza (deterioramenti cognitivi, malattia di Alzheimer, etc.), così come l'assessment psicofisiologico, che si configura come una specifica parte dell'esame psicodiagnostico dedicato alla valutazione delle risposte psicofisiologiche dei soggetti, come la conduttanza cutanea (SCR); la frequenza cardiaca (FC); la pressione sistolica e diastolica; la respirazione; la risposta emozionale, ecc.. Infine, uno degli ambiti di maggior utilizzo dei test è sicuramente quello psicotecnico, per la selezione e valutazione delle competenze lavorative. Anche se viene considerato abitualmente un ambito applicativo proprio della psicologia del lavoro, spesso nei contesti di selezione o di valutazione del potenziale vengono utilizzati specifici strumenti cartacei (questionari, scale, test di livello), colloqui motivazionali e strumenti di valutazione delle capacità gestionali e relazionali basati su forme di role-play, di simulazione di scenari realistici, e di dinamiche di gruppo. Tutte queste tecniche rientrano nel cosiddetto Assessment Center (vedi paragrafo 1.3.1).

Molto spesso gli stessi test sono utilizzati in più ambiti, sebbene con finalità e obiettivi di indagine differenti, oppure sono costruiti specificatamente per la valutazione di determinati aspetti in contesti definiti. In ogni caso, il crescente interesse in tutti questi settori ha determinato nel corso degli ultimi decenni lo sviluppo di un numero enorme di strumenti psicodiagnostici, tanto che da una analisi della letteratura appare difficile individuare un metodo unico per la loro

classificazione, ma piuttosto si rilevano diverse tassonomie centrate principalmente su tre criteri: caratteristiche esterne, modalità di somministrazione e aspetti funzionali. Di seguito una descrizione delle tre classificazioni:

- **Caratteristiche esterne:** i test sono divisi tra test carta e matita, che prevedono una risposta scritta agli item, e test di rendimento, dove l'item è rappresentato da un prestazione manuale, come ad esempio la manipolazione di oggetti.
- **Modalità di somministrazione:** si distingue tra test individuali, sotto la supervisione di un valutatore, e test collettivi, svolti simultaneamente da un gruppo di persone con l'ausilio di istruzioni d'uso.
- **Funzione misurata:** metodo più diffuso di classificazione, si focalizza sull'obiettivo di analisi del test e distingue tra test di personalità, che esplorano tratti della personalità dell'individuo, come gli interessi, l'affettività, la motivazione, ecc.; e i test di efficienza, che si concentrano invece su tutti gli aspetti conoscitivi dell'essere umano come le attitudini, l'intelligenza, le competenze trasversali, il profitto, e così via.

In riferimento alla funzione misurata, è possibile fare in ciascuna tipologia un ulteriore distinzione. Nell'ambito dei test di personalità ad esempio, si possono distinguere test molto diversi a seconda del paradigma personologico di riferimento o al tipo di costrutto misurato o al tipo di valutazione richiesta. In generale, i test di personalità si distinguono in:

1. *Questionari o inventari di personalità:* solitamente test di valutazione piuttosto ampi, basati su un paradigma personologico di ambito cognitivo o della "teoria dei tratti". Sono numerosissimi e sono utilizzati sia nella ricerca che nella clinica. Tra questi i più noti sono l'MMPI, che benché nasca come strumento clinico, il suo utilizzo si è largamente diffuso anche in altri ambiti quali la psicologia del lavoro, la selezione del personale, la psicologia giuridica per scopi peritali, e l'orientamento, e questo grazie al fatto che fornisce sia un profilo normale che patologico di personalità; e il 16PF (16 Personality Factors) di Cattell, che si focalizza proprio sulla misurazione delle caratteristiche normali della personalità umana.

2. *Test obiettivi di personalità*: tecniche derivate dagli esperimenti in laboratorio, che hanno il vantaggio di fornire dati precisi e oggettivi, ma lo svantaggio di esplorare settori limitati della personalità. Attraverso la raccolta di dati fisiologici o all'osservazione diretta in ambito controllato, è possibile investigare caratteristiche come la percezione, lo stress, l'ansia, ecc.. Molti dei test oggettivi oggi in uso sono raccolti nell'Objective-Analytic Test Kit di Cattell e Schuerger.
3. *Test Proiettivi*: sono una tipologia di test di personalità che si basa sulla presentazione di stimoli che i) non sono strutturati o lo sono solo in maniera parziale (macchie di inchiostro, figure umane o animali, ecc), ii) richiedono un processo di interpretazione da parte dei soggetti, iii) non hanno un legame evidente con il costrutto che vanno a misurare, iv) richiedono risposte aperte senza possibilità di errore. Se da una parte gli item dei test proiettivi sono destrutturati e poco chiari, di contro sono caratterizzati invece da un setting di somministrazione rigoso e da modalità di codifica ben definite e standardizzate, dette siglatura. Sono utilizzati soprattutto in contesti psicodinamici, per valutazioni globali della personalità o per l'approfondimento di specifiche tematiche personologiche, relazionali e/o psicopatologiche. Esistono anche in questo caso una vasta gamma di reattivi, tra cui i più noti sono il Test di Rorschach, della metà del '900, finalizzato alla creazione di un profilo delle attitudini e di personalità (Sfera dell'Intelligenza, dell'Affettività e del Contatto Sociale) e per identificare eventuali nodi problematici del soggetto, molto utilizzato per esplorare le dinamiche interpersonali. Il Thematic Apperception Test (TAT) di Murray, ed il Children Apperception Test (CAT), versione per bambini, ossia test tematici focalizzati sul contenuto e gli aspetti fantastici del pensiero espresso dal soggetto nel descrivere l'immagine, e finalizzati all'analisi delle emozioni, degli atteggiamenti e dei processi cognitivi.

Per quanto riguarda invece i test di efficienza, si parla di *test di profitto*, per la valutazione degli esiti della formazione, sia essa scolastica o professionale, come ad esempio il Modern Language Aptitude Test (MLAT) di Carroll e Sapon (1964) o il Logos 10 e Logos 15 di Boncori (1985); *test di attitudine*, per la predizione del successo in certe attività mediante l'identificazione e la valutazione di abilità specifiche, come il Differential Aptitude Tests (DAT) per la determinazione di profili

professionali; e i *test di abilità generale*, per identificare e misurare le abilità del singolo in una specifica area, di cui il più noto è sicuramente il Wechsler Intelligence Scale, per la misurazione dell'intelligenza, sia nella versione per adulti, la WAIS, che per bambini, la WISC.

Di particolare rilevanza ai fini di questo studio, sono proprio i test di efficienza, e in particolare quelli specificatamente costruiti per la valutazione delle competenze trasversali o meglio definite Soft Skill.

1.2 Le Soft Skill

Le Soft skill, o competenze trasversali, rappresentano quel bagaglio di conoscenze, abilità e qualità personali utilizzate nelle relazioni sociali e che ogni individuo arricchisce grazie alle diverse esperienze sperimentate nel corso della propria vita. Con questo termine quindi si fa riferimento ad una serie di qualità e/o di caratteristiche personali che aumentano la conoscenza di se stessi e migliorano le interazioni e le relazioni. La loro espressione è data dalla capacità di un individuo di interagire con gli altri in modo costruttivo, di leggere e comprendere l'ambiente sociale di riferimento, ed utilizzarne le risorse a disposizione in modo coerente ed appropriato, in relazione all'obiettivo.

In letteratura, però, è difficile trovare una definizione universale o onnicomprensiva di soft skill, questo perché si tratta di un concetto ampio che fa riferimento ad ambiti interpersonali e intrapersonali, e che pertanto implica una combinazione di diverse componenti. La difficoltà principale è rappresentata dall'uso indifferenziato di una serie di termini utilizzati spesso come sinonimi, quali abilità, conoscenze, comportamenti, terminologia che si differenziano ulteriormente a seconda del contesto di applicazione, che può essere educativo, psicologico, organizzativo, giuridico, ecc.

Una chiarificazione di questi concetti è stata data nel Quadro Europeo delle Qualifiche e dei Titoli (European Qualification Framework - EQF), presentato nella Raccomandazione del Parlamento Europeo e del Consiglio del 23 aprile 2008. Si tratta di un sistema unitario di certificazione delle competenze dei cittadini, acquisite in ambito formale, non formale o informale considerate competenze essenziali per l'apprendimento nell'intero corso di vita. Secondo l'EQF, ogni cittadino dovrebbe

acquisire queste competenze chiave, combinando conoscenze, abilità e atteggiamenti appropriati al contesto. Viene proposta così una distinzione tra le conoscenze, *definite come il risultato dell'assimilazione di informazioni attraverso l'apprendimento*, e quindi sono rappresentate da quell'insieme di fatti, principi, teorie e pratiche specifiche di un settore lavorativo o educativo, le abilità, *rappresentate dalle capacità di applicare le conoscenze e di utilizzare il proprio know-how per portare a termine compiti e risolvere problemi*, e possono essere sia di tipo cognitivo (uso del pensiero logico, intuitivo e creativo) che pratico (abilità manuale e uso di metodi, materiali, strumenti); e le competenze *date da una comprovata capacità di utilizzare le conoscenze, le abilità e le capacità personali, sociali e/o metodologiche, in situazioni di lavoro o di studio e nello sviluppo professionale e personale*.

In questo quadro, quindi, le conoscenze sono il risultato diretto dell'apprendimento e della formazione in un particolare settore, ma possono essere anche tutte quelle informazioni tratte dalle osservazioni, dalle esperienze, dalle credenze e dai pregiudizi in ogni settore della vita. Si tratta cioè di entità che possono essere archiviate e richiamate alla memoria. Indipendentemente da come è acquisita, la caratteristica principale della conoscenza è che può essere facilmente verificata, ad esempio attraverso un esame. In questo caso l'apprendimento è verificato e valutato in termini comportamentali. E anche se *il passare del tempo rende sempre più difficile richiamare alla memoria saperi astratti e quanto più tempo trascorre tanto più diventa difficile recuperarli* (Baddeley 2001), la valutazione comportamentale della conoscenza rimane comunque la soluzione più adatta, perché altrimenti non sarebbe possibile verificare nessun sapere, dato che il sistema cognitivo umano non è in grado di conservare conoscenze in maniera definitiva.

Per gestire le conoscenze, l'essere umano utilizza le abilità o skill, ossia attività mentali che si verificano nel cervello mentre utilizziamo, trasformiamo o potenziamo le conoscenze a disposizione. *Le skill sono spesso associate con attività cognitive di alto livello come il problem solving, il ragionamento, il pensiero, la capacità di trarre conclusioni ed includono le capacità di analisi, sintesi e valutazione finalizzate a riprodurre o ad estendere le strutture cognitive esistenti* (Bloom 1984). Pur essendo per definizione processi mentali interni, le abilità cognitive possono esprimersi in comportamenti osservabili quando si interagisce con la realtà. Più nello

specifico si verifica un doppio processo, da una parte l'individuo modifica l'ambiente circostante ogni qualvolta utilizza le proprie conoscenze per compiere azioni pianificate in un dato contesto, dall'altra, le continue stimolazioni dell'ambiente esterno producono nuove esperienze, nuove conoscenze e, pertanto, un continuo adattamento e modellamento delle strutture cognitive interne, che vanno così a migliorare la performance, fino a far diventare talune abilità totalmente automatizzate e inconsapevoli. Ed è proprio l'automatizzazione il principio utilizzato nella formazione e valutazione delle skill cognitive sebbene, per la sua stessa natura, le rende poi difficilmente distinguibili dai comportamenti, determinando l'impossibilità di effettuare valutazioni certe. Pertanto, non potendo valutare i processi mentali, la padronanza di abilità specifiche è di solito testata in termini operazionali-comportamentali, attraverso cioè i comportamenti collegati direttamente al loro utilizzo. Tuttavia la sequenza pensiero-azione, seppur presente nelle fasi iniziali di apprendimento, sparisce nel momento in cui la performance si automatizza. Per questo motivo, molte tecniche utilizzano prima l'esperienza e solo successivamente procedono all'acquisizione di abilità cognitive. In altri casi invece questa relazione assume un'espressione simultanea, come ad esempio il linguaggio, inteso come abilità di costruire frasi dotate di senso che siano immediatamente operazionalizzate. Questo stretto legame determina grande confusione sul significato specifico dei due termini, tanto che Chomsky, nel 1960, distingue tra l'atto del parlare (performance linguistica), ossia espressione individuale legata alle prestazioni, e abilità cognitive ad essa associata (competenze linguistiche), che sono invece generali ed invariabili, gettando le basi per gli studi sul concetto di competenza.

A partire dalla visione funzionalista di Chomsky, sono stati numerosissimi gli autori che hanno fornito una definizione di competenza, spesso con sfumature di significato strettamente connesse all'approccio teorico di riferimento o all'ambito di applicazione. La Tabella 1 mostra solo alcune tra le principali definizioni di "competenza" proposta negli ultimi decenni.

Una competenza relativa ad una mansione è una caratteristica intrinseca di un individuo che porta ad una performance efficace o superiore nella mansione (Boyatzis, 1982)
La competenza è la capacità di dar luogo ad una sequenza regolare di comportamento coordinato, efficace rispetto agli obiettivi, dato il contesto in cui ha luogo” (Nelson Winter, 1982)
La competenza è il bagaglio di capacità che la persona possiede e alle quali fa ricorso per rispondere a specifiche richieste della posizione lavorativa (Ratti, 1989)
Capacità di reinvestire le proprie acquisizioni precedenti e di combinarle per risolvere nuovi problemi. E’ il sapere reinvestito nell’azione, i saperi in atto (Malglaiive, 1990)
Saper mobilitare le proprie conoscenze e qualità per far fronte a un problema dato, le competenze definiscono le conoscenze e le qualità messe in atto in diverse situazioni (Mandon, 1990)
Le competenze sono caratteristiche individuali intrinseche che indicano i modi di comportarsi o di pensare che si ripetono nelle loro grandi linee nelle diverse situazioni e perdurano per un periodo di tempo ragionevolmente lungo (Guion, 1991)
Una competenza può essere definita come una caratteristica intrinseca individuale, causalmente collegata a una performance efficace o superiore in una mansione o in una situazione che è misurata sulla base di un criterio prestabilito (Spencer & Spencer, 1993)
La competenza non è uno stato o una conoscenza posseduta. Non è riducibile né a un sapere, né a ciò che si è acquisito (...). La competenza non risiede nelle risorse (conoscenze, capacità cognitive, relazionali ...) da mobilitare, ma nella mobilitazione stessa dei saperi che si sono saputi selezionare, integrare e combinare in un contesto e per un obiettivo specifico. La competenza rientra nel concetto di “saper mobilitare”. La competenza è un “saper agire” riconosciuto. Qualunque competenza, per esistere, necessita del giudizio altrui” (Le Boterf, 1995)
La competenza è concepita per lo più come ‘prestazione’, ma si deve piuttosto intendere come mobilitazione e orchestrazione di risorse cognitive, affettive e operative che permettono di generare un’infinità di pratiche adatte a situazioni sempre rinnovate; in questo modo conoscenze, abilità e atteggiamenti’ interagiscono e si ricompongono nel concetto olistico di ‘competenza’ supportata dalla interazione sociale (Michele Pellerey, 2002)
Patrimonio complesso di conoscenze, abilità e risorse individuali di cui un soggetto dispone per realizzare il proprio sviluppo personale, per affrontare in modo efficace l’inserimento socioculturale e lavorativo (Invalsi, 2004)

Tabella 1: Principali definizioni di Competenza

1.2.1 Competenza vs Competenza Trasversale

Un’attenta analisi delle definizioni mostrate nella Tabella 1, evidenzia che sebbene all’inizio del secolo, il comportamentismo definiva la competenza come un comportamento osservabile e misurabile, risultato di uno schema rigido e funzionale

al compito e non legata all'individualità e alla specificità di un soggetto, a partire dalla seconda metà del Novecento, i cognitivisti e ancor più i costruttivisti, hanno cominciato a delineare la mente come qualcosa di dinamico e creativo, dando un posto centrale alle operazioni cognitive (motivazione, attenzione, osservazione, percezione, memorizzazione, metacognizione) capaci di strutturare l'esperienza e l'apprendimento e costruire competenze, attestabili in singole prestazioni. Le competenze non sono più quindi intese come un saper fare "rigido e precostituito" dagli ambiti operativi, ma come un patrimonio dinamico che il soggetto acquisisce e sviluppa attraverso esperienze diverse, occasionali ed intenzionali, e che gli permette quindi di partecipare in maniera creativa ed originale ad un ampio numero di attività.

Ed è in questo quadro che si approfondisce lo studio delle competenze distinguendo innanzitutto tra **Soft Skill** (competenze trasversali) e **Hard Skill** (competenze tecniche). Mentre con il termine hard, si fa riferimento al ventaglio delle conoscenze tecniche e professionali, acquisite solitamente attraverso programmi di formazione, per lo svolgimento di una specifica mansione all'interno di un certo dominio o settore, molto più complesso risulta il tentativo di decidere cosa includere o escludere nella definizione di soft skill, sebbene si evidenzia un generale accordo nel ritenerle idealmente complementari alle hard skill e visibili in tutti i settori di vita quotidiana, e non solo quelli professionali. Il tentativo di identificare cosa sono e quali sono le soft skill, ha portato alla nascita di studi che hanno ampliato il concetto di competenza agganciandolo a quello di intelligenza, come ad esempio, Gardner (1983) che introduce il termine di intelligenza multipla, o Marlowe (1986) che parla invece di intelligenza sociale o Daniel Goleman (1995), che invece definisce il concetto di intelligenza emotiva.

Howard Gardner nel suo famoso libro *Formae mentis* del 1983 non solo mette in crisi il concetto di Intelligenza⁴, affermando che non si tratta un'entità statica, geneticamente ereditata e non potenziabile nel corso della vita, ma anche tutti i test

⁴ Charles Spearman, noto comportamentista, all'inizio del novecento definì l'intelligenza con il termine Fattore G, identificabile con una capacità comune e misurabile in tutti gli individui.

sul quoziente intellettivo⁵, misura di questa unica forma di intelligenza (vedi paragrafo 1.1.1) sviluppati fino a quel momento. Secondo Gardner, ogni individuo è caratterizzato da una molteplicità di intelligenze, ovvero potenzialità biologiche presenti sin dalla nascita che assumono una particolare combinazione di livelli di sviluppo in ogni essere umano, rendendo il suo profilo intellettivo del tutto unico. *L'intelligenza multipla è vista come insieme di capacità e competenze che tutti gli individui possiedono, nessuna più importante delle altre, ma con livelli e capacità combinatorie differenti, che rende ogni persona intelligente a modo proprio.*

Nella sua prima elaborazione, Gardner distingue sette diverse intelligenze: 1. Logico-matematica 2. Linguistica 3. Spaziale 4. Musicale 5. Cinestetica 6. Intrapersonale 7. Interpersonale. A queste, nei suoi scritti successivi (1999)⁶ aggiunge altre due forme: Naturalistica ed Esistenziale. Le 9 intelligenze sono descritte dettagliatamente nella tabella che segue.

INTELLIGENZE	DEFINIZIONI
Intelligenza linguistica	Permette agli individui di comunicare e di costruire il significato del mondo attraverso il linguaggio. L'espressione più alta è offerta dai poeti, scrittori, linguisti, filologi, oratori. Più in generale, è spiccata nelle persone che amano giocare con le rime, che fanno giochi di parole, che hanno sempre una storia da raccontare, che acquisiscono velocemente altri linguaggi
Intelligenza musicale	Permette alle persone di creare, comunicare e comprendere i significati dei suoni. Tipica dei compositori e musicisti, la si nota in quei soggetti attenti ai suoni del mondo o che tamburellano continuamente ritmi complicati con le dita.
Intelligenza logico-matematica	Permette agli individui di usare e di apprezzare le relazioni astratte, e caratterizza in particolare scienziati, matematici e filosofi. Più genericamente, è visibile nei soggetti analitici, che prima esaminano attentamente le parti di un problema, e poi di verificano sistematicamente le soluzioni.
Intelligenza spaziale	Rende possibile alle persone di percepire informazioni visive o spaziali, di trasformare tale informazione e di ricreare immagini visive tratte dalla memoria. Presenti ad un livello alto negli architetti, scultori e ingegneri, sono visibili anche nei soggetti che amano le mappe concettuali, grafici e tabelle.

⁵ Gardner (1983) descrive gli studi sul QI come un movimento empirico cieco, che basa la valutazione dell'intelligenza sulla capacità di dare la risposta corretta, in termini di risultato osservabile e non nei termini di processo di pensiero sottostante.

⁶ Intelligence Reframed: Multiple Intelligence for the 21st Century, Howard Gardner, Basic Books, 1999

Intelligenza corporeo-cenestetica	Permette alle persone di usare il proprio corpo, per creare prodotti o risolvere problemi, come nel caso di atleti, chirurghi, danzatori, coreografi e artigiani. Evidente in tutti quei soggetti che praticano uno sport o preferiscono i lavori manuali.
Intelligenza Intrapersonale	Permette alle persone di comprendere la propria individualità, di saperla inserire nel contesto sociale per ottenere risultati migliori nella vita personale, e anche di sapersi immedesimare in ruoli e sentimenti diversi dai propri. Sebbene è presente in tutti, anche con livelli diversi, è particolarmente spiccata negli attori.
Intelligenza interpersonale	Permette alle persone di comprendere gli altri, le loro esigenze, le paure, i desideri nascosti, di creare situazioni sociali favorevoli e di promuovere modelli sociali e personali vantaggiosi. La sua espressione massima è visibile nei politici, leader, imprenditori di successo, psicologi, ma è visibile anche in tutti quei soggetti impegnati in attività di volontariato o attività di sostegno agli altri.
Intelligenza Naturalistica	Permette di individuare determinati oggetti naturali, classificarli in un ordine preciso e cogliere le relazioni tra di essi. È l'intelligenza tipica di biologi, astronomi, antropologi, medici e altri.
Intelligenza Esistenziale	Permette di riflettere consapevolmente sui grandi temi dell'esistenza, come la natura dell'uomo, e di ricavare da sofisticati processi di astrazione delle categorie concettuali che possano essere valide universalmente. È tipica dei filosofi e degli psicologi, e in parte anche dei fisici.

Tabella 2: Le intelligenze multiple di Gardner

La sua classificazione nasce da un lavoro di rassegna di un numero enorme di ricerche in diversi ambiti, quali biologia, antropologia, psicologia cognitiva e dello sviluppo, neuropsicologia e psicomètria, che gli ha permesso di individuare 8 criteri (la lista completa è mostrata nella Tabella 3) che, secondo l'autore, permettono di identificare le specifiche intelligenze. Tra questi, particolarmente rilevante è l'identificazione delle “operazioni centrali”, ossia quell'insieme di processi dell'informazione o meccanismi che si occupano di uno specifico genere di input.

1	La sua potenziale isolabilità per un danno al cervello
2	La sua presenza in soggetti speciali come gli idiots savants, i bambini con deficit, le persone prodigio o altri individui eccezionali
3	La sua connessione ad un identificabile gruppo centrale di operazioni che riguardano i processi dell'informazione o meccanismi che si occupano di uno specifico genere di input
4	Una distintiva storia del modo in cui si è sviluppata tale abilità insieme con la definizione dei comportamenti che ne definiscono lo stato finale
5	Una storia della sua evoluzione e una plausibilità della sua presenza
6	Essere supportata da compiti sperimentali e psicologici
7	Essere supportata da ricerche psicomètriche
8	Possesso di una predisposizione genetica a cristallizzarsi in un sistema simbolico.

Tabella 3: Lista dei criteri di Gardner per l'identificazione delle intelligenze

Sinteticamente, secondo l'autore, ogni intelligenza è l'insieme di una serie di capacità che se da una parte permettono la risoluzione di problemi, dall'altra sono in grado di trovare o creare problemi, al fine di aumentare o potenziare le conoscenze.

Delle nove intelligenze identificate da Gardner, due sono particolarmente interessanti in questa sede, perché strettamente connesse al concetto di soft skill: **l'intelligenza intra e inter personale**. La combinazione tra la prima, riferita alla capacità di capire se stessi, le proprie peculiarità e potenzialità, e la seconda, inerente invece alla comprensione degli altri, permettono la creazione di rapporti e relazioni, l'acquisizione di maggiori conoscenze sia di tipo personale che sociale, e lo sviluppo di una serie di competenze trasversali, quali la comunicazione (verbale e non verbale), l'empatia, la socialità, la mediazione. Sulla stessa scia, si inseriscono i lavori sull'intelligenza sociale ed emotiva di Marlowe. Quest'autore, infatti, basandosi sugli studi pionieristici sull'intelligenza sociale di Thorndike (1920), definisce *l'intelligenza sociale come quell'insieme di competenze trasversali che permettono all'individuo di agire in modo appropriato in ogni situazione sociale attraverso la comprensione dei sentimenti, pensieri e comportamenti degli individui (compresi se stessi), che gli permettono così di sperimentare significative esperienze affettive* (Marlowe, 1986). Nella sua teorizzazione, l'autore classifica queste competenze trasversali in cinque dimensioni principali: atteggiamento prosociale (interesse e preoccupazione per gli altri), le capacità di performance sociale (adeguata interazione con gli altri), capacità empatica (capacità di identificarsi con gli altri), espressività emotiva (modo di esprimere le emozioni verso gli altri), fiducia in sé e negli altri. Molto vicino a queste idee è la proposta di Goleman (1995) di intelligenza emotiva. Per questo autore, le soft skill vengono definite come quell'insieme di capacità personali, quali la creazione di rapporti costruttivi, l'intuito, l'analisi delle relazioni con gli altri, che combinati con elementi di personalità quali l'ottimismo, l'adattabilità, lo spirito di iniziativa, vanno a formare la base per mantenere e accrescere il proprio successo, in ogni situazione e ambiente. Goleman inoltre (1998), restringendo l'analisi ad un ambito strettamente professionale, afferma che il successo è fortemente connesso alle capacità dell'individuo di gestire i sentimenti, interagire e comunicare con gli altri, piuttosto che a punteggi statici del quoziente di intelligenza (QI) o alla manifestazione di sole competenze tecniche.

Una buona fiducia in se stessi, la piena consapevolezza di sé, l'autocontrollo, l'impegno e l'integrità sono, secondo l'autore, i fattori che distinguono i dipendenti di successo, quegli individui cioè che permettono la creazione di un business aziendale maggiormente efficace. Questi uomini e donne, questi leader, sono caratterizzati da *una spiccata intelligenza emotiva, che include la capacità di motivare e influenzare gli altri, fornire feedback efficaci, costruire buone relazioni, monitorare il proprio comportamento, gestire le emozioni, sia di sé che degli altri, e leggere le situazioni interpersonali.*

Quello che emerge dalla breve analisi di queste definizioni e teorie, è che una delle caratteristiche principali che rende le **soft skill** un tema particolarmente interessante è che esse **possono essere arricchite nel corso della vita**, grazie alle esperienze personali e professionali, **e quindi di fatto possono essere valutate, migliorate e apprese**; non solo in maniera inconsapevole e indiretta, ma in maniera diretta e cosciente, attraverso opportuni sistemi di formazione e valutazione. Pertanto accanto allo sviluppo di teorie e definizioni, l'interesse si è contestualmente focalizzato, soprattutto in ambito organizzativo, sui metodi e metodologie per valutare queste competenze, finalizzate allo sviluppo di percorsi formativi personalizzati e mirati al loro miglioramento e potenziamento.

1.2.2 I modelli delle competenze

L'analisi della letteratura evidenzia in particolare due modelli principali finalizzati a valutare le soft skill: il **modello americano** e il **modello inglese**.

Il padre dell'approccio americano può essere considerato David McClelland che in un suo studio nel 1973, sulla selezione di diplomatici americani, affermò che la *“cultura scolastica ed i titoli di studio non sarebbero stati più sufficienti per predire il successo professionale”*, ma che questo è dipendente da caratteristiche personali legate alla motivazione e alla realizzazione. Con il suo lavoro, e i test da lui realizzati, McClelland apre così la strada allo studio delle competenze come elementi essenziali per il successo lavorativo, diventando un modello consolidato di individuazione e definizione delle soft skill, e punto di riferimento per molte altre ricerche in questo settore. La sua definizione viene ulteriormente approfondita da Boyatzis nel 1982 (vedi Tabella 1) che ha sua volta viene ripresa da Spencer e

Spencer nel 1993. Questi ultimi propongono una definizione composta di due parti: la prima ne individua il significato e cioè «una caratteristica intrinseca di un individuo causalmente collegata ad una performance eccellente in una mansione», la seconda ne elenca invece i suoi tratti costitutivi, essa «si compone di motivazioni, tratti, immagine di sé, conoscenze, e abilità». I due autori propongono la metafora dell'Iceberg per rappresentare il loro modello (Figura 2) , dove le conoscenze e le abilità sono rappresentate nella parte alta dell'iceberg, in quanto relativamente osservabili; mentre l'immagine di sé, i tratti e le motivazioni sono, invece, “sommersi”, cioè nascoste nelle parti più intime della personalità.

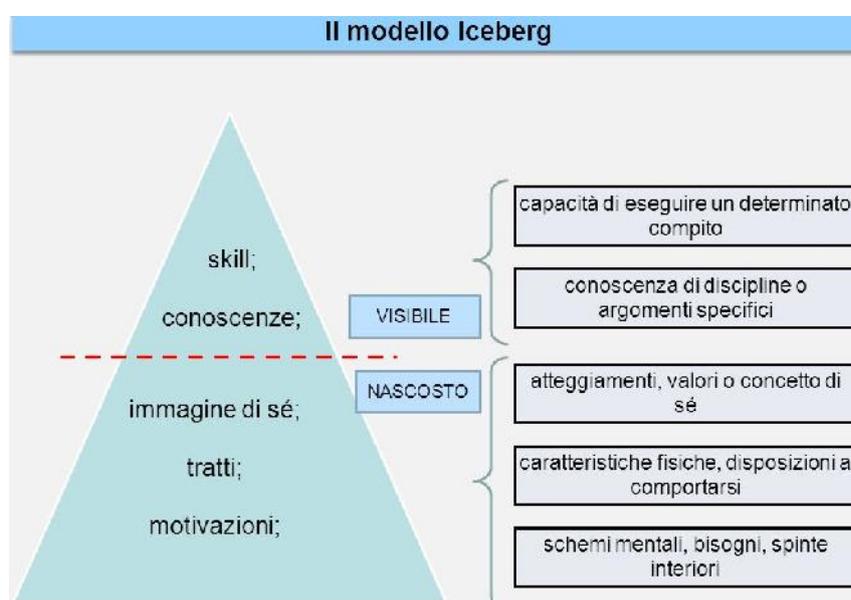


Figura 2: il modello di Spencer & Spencer

In un contesto di apprendimento, sia questo professionale o non, le motivazioni, i tratti e l'immagine di sé sono i fattori più semplici da identificare in una fase di selezione, ma risultano i più difficili da trattare e potenziare in un processo di formazione. Tuttavia, continuano gli autori, è solo sulla base di questi che si può effettuare una buona predizione proprio delle skill di comportamento/azione, che a loro volta sono predittive dei risultati della performance nella mansione. Secondo gli autori, la diversa combinazione di queste componenti, permette di distinguere tra:

i) **Performance efficace:** cioè un livello di lavoro “minimo accettabile”, la soglia al di sotto della quale una persona non può essere considerata competente nel ruolo;

ii) **Performance superiore**: definita statisticamente come una deviazione standard dalla performance media, o grosso modo il livello raggiunto da una persona su dieci, in una situazione di lavoro data.

Il modello americano è spesso messo a confronto con quello inglese, sostenuto dal National Vocational Qualification⁷, che valuta le competenze individuali osservando la persona mentre svolge le principali mansioni lavorative, verificando concretamente se è in grado di svolgerle in maniera ottimale. La competenza in questo quadro è definita come “*un determinato compito lavorativo che la persona è in grado di svolgere secondo un livello predefinito*”. Questo significa che le competenze di un individuo e i compiti in grado di svolgere in un determinato ambito coincidono. Si parla pertanto di standard di prestazione (National Occupational Standards⁸), ossia l’insieme di quelle abilità necessarie per essere considerati competenti in quel ruolo o mansione. Questi standard permettono, nel Regno Unito, di ottenere qualifiche per un gran numero di professioni, indipendentemente dal tipo di apprendimento realizzato fino a quel momento. In altre parole questo approccio, si focalizza sull’apprendimento non formale, in quanto mette al centro della valutazione delle competenze ciò che la persona sa fare e non come lo ha imparato.

Il modello inglese può essere più affidabile per la valutazione delle immediate prestazioni lavorative, ma non fornisce garanzie rispetto all’affidabilità del soggetto in contesti variabili e complessi, aspetto su cui invece si focalizza il modello americano. Per questo motivo, oltre ad essere utilizzato nei contesti professionali, il modello americano è particolarmente idoneo in contesti formativi con finalità di orientamento. In Europa, questo modello è usato sia, in Italia, dall’ISFOL⁹, Istituto per lo Sviluppo della Formazione Professionale dei Lavoratori, che dai Centri Inter-Istituzionali per il Bilancio delle Competenze¹⁰, in Francia. Entrambi questi enti promuovono protocolli finalizzati a verificare se la persona possiede quelle competenze da cui dipende una prestazione lavorativa di buon livello. Ci si focalizza

⁷ <http://www.naric.org.uk/naric/>

⁸ <http://www.ukstandards.org.uk/Pages/index.aspx>

⁹ <http://www.isfol.it>

¹⁰ <http://www.cibc.net/>

su un insieme di competenze trasversali e non, che vengono rintracciate tramite colloqui e test, con l'obiettivo di personalizzare la formazione dei lavoratori e migliorare così la loro performance.

Questi modelli hanno permesso di fare un po' di chiarezza e di identificare quelle che possono essere considerate le soft skill principali. Ad oggi, infatti, dall'analisi della letteratura è possibile estrapolare 14 soft skill considerate fondamentali per lo sviluppo e la crescita personale e professionale, che sono:

1. Autonomia;
2. Problem solving;
3. Adattabilità;
4. Team working;
5. Intraprendenza;
6. Creatività;
7. Senso di realtà;
8. Assertività/autorevolezza;
9. Precisione;
10. Emotività;
11. Decision making;
12. Adeguatezza al ruolo;
13. Comunicazione verbale e non;
14. Negoziazione.

Di queste, la comunicazione verbale e non verbale, ma ancor di più la negoziazione, sono le soft skill su cui si concentra il presente lavoro. Per la valutazione di queste competenze sono stati creati diversi test, spesso inseriti in scale più ampie, e quasi sempre utilizzati nell'ambito professionale con obiettivi diversi, dalla selezione del personale, alla valutazione del potenziale o alla pianificazione di percorsi formativi personalizzati. Alcuni esempi di particolare interesse sono:

- **ROCI I e ROCI II di Rahim et al.**, il primo pensato per misurare le dimensioni del conflitto all'interno delle organizzazioni, il secondo per misurare i modi di affrontare il conflitto interpersonale. **Quest'ultimo test rappresenta il punto di partenza dell'intero lavoro qui presentato ed è descritto in maniera approfondita nel capitolo 2.**
- **Questionario stili di comunicazione/negoziazione di Thomas Kilmann:** specifico per valutare gli stili preferenziali di comunicazione e negoziazione.
- **Fifteen Factor Questionnaire (15FQ+)**, basato sul lavoro di Cattell, il test valuta lo stile di leadership, lo stile di collaborazione con i superiori, lo stile relazionale adottato all'interno dei gruppi di lavoro, l'attitudine alla vendita e lo stile di persuasione.

Tutti questi test sono normati e standardizzati in moltissimi paesi e quindi disponibili in molte lingue. Ad oggi, possiamo affermare che sono tra i più utilizzati per

misurare queste competenze. Come quasi la maggior parte dei test, anche questi rientrano nella categoria dei test “carta e matita” (vedi paragrafo 1.1.3), sebbene le recenti innovazioni tecnologiche, le esigenze specifiche dei campi di applicazione e il sempre maggior interesse verso le possibilità offerte dal web e dall’era digitale, stanno determinando un graduale passaggio verso modalità più tecnologiche, che sfruttano i principi della partecipazione a distanza, della gamification e della realtà virtuale in 3D, dapprima per scopi formativi e, più recentemente, per fini valutativi.

1.2.3 Le soft skill e le nuove tecnologie

Come abbiamo visto, le soft skill o competenze trasversali hanno destato un ampio interesse sia perché ritenute fondamentali per la crescita personale e collettiva in qualsiasi settore, privato e professionale, ma soprattutto perché esse, una volta valutate, possono essere migliorate e potenziate, grazie ad opportuni percorsi formativi. In questo quadro, risulta molto forte lo sforzo di molti settori, da quelli accademici e formativi, a quelli organizzativi e istituzionali, nel creare protocolli e metodologie di valutazione e formazione quanto più funzionali all’obiettivo. Il contestuale e costante sviluppo delle nuove tecnologie, proprio degli ultimi anni, ha fatto sì che questa ricerca utilizzasse sempre più prodotti e strumenti tecnologici, con il risultato di un passaggio da metodi “tradizionali”, quali la formazione in aula o la valutazione “carta e matita”, a metodi digitali, trasformando di fatto il mondo della formazione e, seppur ancora in maniera limitata, quella della valutazione delle competenze trasversali.

È in questo quadro che relativamente poco tempo fa è nato l’e-learning, ossia l’uso della “ *tecnologia per progettare, distribuire, selezionare, amministrare, supportare e diffondere la formazione*”, come è stato definito la prima volta da Elliott Masie esperto americano di tecnologie. Più nel dettaglio, l’e-learning è una metodologia di insegnamento e apprendimento che coinvolge sia il prodotto, rappresentato da qualsiasi materiale o contenuto messo a disposizione in formato digitale attraverso supporti informatici o di rete, sia il processo formativo, ossia la gestione dell’intero iter didattico che coinvolge tutti gli aspetti di un percorso di formazione, dall’erogazione, alla fruizione, all’interazione, alla valutazione. I principali vantaggi di questa modalità di insegnamento sono rappresentati dai servizi di assistenza e di

tutoring, dall'interazione sincrona e asincrona, dalla possibilità di condividere e collaborare a livello di community. Ma più di tutto, la flessibilità dei sistemi di e-learning offrono la possibilità agli utenti di reperire in maniera facile ed economica i contenuti formativi, consentendo così la creazione di percorsi autogestiti di formazione. Questa facilità di fruizione cresce parallelamente alle innovazioni e agli sviluppi delle tecnologie. Così la sempre maggiore diffusione di notebook, smartphone e tablet e il conseguente uso sempre più massiccio della rete internet, in ogni luogo e in ogni dove, ha modificato totalmente lo stile di vita delle persone, sia a livello personale che professionale, dando a tutti la possibilità di usufruire di servizi di apprendimento online sempre più semplici e intuiti, secondo la logica, cara all'e-learning, del learning by doing (vedi paragrafo 1.3.1), ossia imparare facendo. Nell'ambito della formazione delle soft skill, questa trasformazione è stata ancora più straordinaria, grazie alle caratteristiche proprio delle competenze trasversali, che bene si sposano con metodi non tradizionali di formazione, essendo come abbiamo detto un ventaglio di abilità che si apprendono nel corso della vita in situazioni formali, informali e non formali. Così la formazione e ancor di più la formazione delle soft skill, avvalendosi delle innovazioni tecnologiche, e in particolare dei sempre maggiori miglioramenti nella creazione di ambienti virtuali e delle relative grafiche, entra nel campo dei video giochi. Questo fenomeno è reso ancor più incisivo dagli sforzi della Comunità Europea, che negli ultimissimi anni, ha finanziato un numero elevato di progetti finalizzati allo studio del connubio tra la tecnologia propria dei videogiochi e gli strumenti diagnostici/formativi. Questo emergente settore di ricerca e sviluppo industriale nasce dall'intersezione di molteplici discipline quali la psicologia, l'ingegneria del software, le tecnologie del videogaming (detta gamification), la pedagogia e l'organizzazione aziendale ed è noto come **Serious Gaming** (vedi paragrafo 1.3). Da un punto di vista metodologico, esso non è altro che l'estensione di tradizionali pratiche educative/formative (come per esempio il role-play, i simulation games e le esercitazioni guidate nei laboratori) basate sul coinvolgimento attivo dei discenti e dei docenti in un ambiente digitale rappresentato, per l'appunto, come un videogioco. In breve, la prospettiva di coniugare l'efficienza delle tradizionali e consolidate pratiche di intervento e valutazione psicopedagogica con l'estrema flessibilità di utilizzo e distribuzione degli

attuali prodotti digitali rappresenta l'ambizione dell'approccio dei Serious Games che ne sta facendo uno delle nicchie di innovazione di prodotto/servizio a più alto tasso di crescita nell'ambito della gestione delle risorse umane.

1.3 I Serious Game

Il termine Serious Games (SGS) risale agli inizi degli anni 70, grazie ad un lavoro di Abt, sebbene si è cominciato a discuterne solo nel 2002, grazie a B. Sawyer e D. Rejeski e all'iniziativa Serious Games, ossia l'utilizzo di giochi per la formazione di particolari tipologie di lavori, come ad esempio la formazione del personale dell'esercito. Da allora si è acceso un intenso e interessante dibattito, che ha riguardato innanzitutto la definizione di SGS.

Attualmente la definizione maggiormente accettata è quella proposta da Michael and Chen nel 2006, secondo la quale i *“Serious Games (SGS) sono tutti quei giochi progettati per un obiettivo primario diverso dal puro intrattenimento, con elementi educativi/formativi, realizzati per l'industria quali istruzione, salute, gestione delle emergenze, ingegneria, forze armate, ecc.”*

Di fatto i SGS altro non sono che simulazioni di eventi reali o processi, progettati allo scopo di risolvere un problema. Non si focalizzano sul divertimento e il loro scopo principale è quello di formare ed educare gli utenti, anche se può avere altre finalità, come il marketing o la pubblicità. Non sono un genere di gioco, ma una categoria di giochi con scopi diversi, come ad es. giochi educativi e manageriali, giochi politici, e così via.

La definizione puntuale di SGS ha portato ad interrogarsi sulle differenze tra «SERIOUS» (SGS) e «GAMES» (Giochi di intrattenimento), e in particolare su due domande principali: 1. Cosa rende un gioco «serious»? e soprattutto 2. I videogiochi possono essere usati con scopi educativi?

Il primo a dare una risposta a questi interrogativi fu Jenkins, nel 2009, il quale utilizza per la prima volta il termine Serious Gaming, indicando un'area più vasta che include oltre ai SGS, qualsiasi videogioco usato per scopi educativi, anche se tale dimensione non è necessariamente insita nel software. In altre parole, l'autore utilizza l'aggettivo “serious” anche per tutti quei video giochi nati per motivi

prettamente ludici che in un secondo momento sono stati utilizzati per scopi formativi con o senza modifiche da parte degli sviluppatori (Figura 3).



Figura 3: Rappresentazione dei Serious Gaming

Riguardo agli aspetti pedagogici, i SGS trovano il loro fondamento teorico nelle teorie dell'apprendimento costruttivista, secondo la quale la conoscenza è creata attraverso l'esperienza che si costruisce esplorando il mondo ed eseguendo attività. Di fatto essi altro non sono che ambienti virtuali, in genere in 3D, in cui il giocatore può acquisire conoscenza attraverso l'esplorazione di un ambiente e la pratica di alcune attività precedentemente definite (come ad esempio, la manipolazione degli oggetti), individualmente o eventualmente in collaborazione con altre persone.

Attualmente, nella ideazione e implementazione di SGS si utilizzano vari modelli pedagogici, tra cui in particolare il modello dell'apprendimento esperienziale. Si ritiene infatti che anche le esperienze virtuali, così come quelle reali, sono caratterizzate dall'acquisizione di informazioni e sensazioni, che restano fortemente impresse e permettono al giocatore di affinare percezione, attenzione e memoria favorendo modifiche comportamentali attraverso il **learning by doing** (imparare facendo). *L'apprendimento non si concretizza più in un mero trasferimento passivo di conoscenze tra un docente e un discente, ma in un'acquisizione attiva e consapevole di nuove conoscenze, attraverso una sperimentazione continua e costante, supervisionata e guidata da un esperto.* Questi elementi rendono così l'apprendimento un'esperienza interessante e anche "divertente" che accresce la motivazione e la volontà degli utenti a completare il proprio percorso formativo.

Una fotografia del fenomeno di diffusione dei SGS è mostrata da un'analisi del mercato internazionale che evidenzia un aumento continuo sia nella produzione che nell'utilizzo di questi prodotti, che sembra aver invaso ormai un gruppo variegato di target di utenza. Analizzando la letteratura e muovendosi negli ambienti di gioco online, si possono identificare alcuni motivi specifici di questa vasta popolarità. In particolare, i serious game:

- 1) utilizzano le più recenti tecnologie di simulazione e visualizzazione, che permette al giocatore di “immergersi” in ambienti realistici e sfidanti;
- 2) propongono nuove modalità di interazione (ad esempio, la collaborazione online, i comandi verbali, controllo gestuale, ambienti sociali, giochi di famiglia) aumentando la motivazione e il divertimento;
- 3) supportano la sperimentazione libera dei giocatori che integrano così l'apprendimento formale, e incoraggiano gli studenti a esplorare le varie situazioni, con barriere limitate di spazio e tempo;
- 4) sono economici, soprattutto quelli che forniscono simulazioni di ambienti complessi o costosi e situazioni critiche o di pericolo;
- 5) forniscono modalità multiplayer, favorendo il team-building, la collaborazione e la cooperazione nell'affrontare le sfide;
- 6) possono essere giocati in molteplici situazioni e con tutti i tipi di device, compresi i dispositivi mobili (es. smartphone, tablet), aprendo così nuove prospettive anche per l'apprendimento e la socializzazione on-line;
- 7) propongono nuove tipologie di giochi (ad esempio sfide intellettuali, giochi sociali) che attraggono una popolazione sempre più vasta ed eterogenea, che non si limita quindi ai soli power-gamers, ossia giovani maschi appassionati di esperienze immersive in 3D, ma anche un pubblico femminile o bambini.

In generale, quello che emerge è una visione positiva rispetto a questi prodotti. In altre parole i SGS sono considerati «ambienti formativi di fatto».

Tuttavia, nonostante la grande diffusione e interesse, forti sono ancora i dubbi relativamente all'efficacia e l'efficienza dei SGS. In altre parole, gli esperti si chiedono se con i SGS si impara davvero. Sebbene non ci siano molti studi

“metodici” e completi a sostegno, si riscontra in letteratura un generale consenso riguardo al potenziale educativo dei SGS. Tra le rassegne della letteratura più ampie possiamo annoverare:

- 1) Livingston et al. (1973), che ha analizzato 150 studi in 7 anni, e Chin et al. che hanno replicato l’analisi nel 2009 giungendo a risultati simili, ossia che i “giochi di simulazione” sono in grado di insegnare le informazioni fattuali, anche se non sono necessariamente più efficaci di altri metodi di istruzione.
- 2) Connolly et al. (2012) ha redatto uno studio basato su 129 report, secondo il quale i SGS influenzano i giocatori sul piano percettivo, cognitivo, comportamentale, affettivo e motivazionale, che si traduce in risultati in termini di apprendimento, sebbene questi risultati non sempre sono duraturi nel tempo.

Quello che è stato indagato in maniera più approfondita è l’efficacia dei SGS in contesti specifici e peculiari, dove sembrano ottenere risultati migliori e più solidi. Una serie di studi scientifici hanno infatti mostrato un buon riscontro di questi strumenti innanzitutto nell’assistenza sanitaria, come ad es. per il trattamento di fobie, delle ustioni o trattamenti chemioterapici, misurati attraverso l’uso della Functional Magnetic Resonance Imaging (fMRI). Cole, nel 2012, conclude la sua analisi affermando che il gameplay sembra essere un ingrediente chiave per influenzare il comportamento di salute positivo. Ottimi risultati sono stati ottenuti anche per il miglioramento di alcuni atteggiamenti, come ad es. l’autostima (vedi The Matrix), o per l’acquisizione di conoscenze, come mostrato da Bellotti et al. nel 2012 con uno studio sulle nozioni di beni culturali.

Sebbene le evidenze scientifiche a sostegno della validità di questi test sono esigue, bisogna sottolineare che nella maggior parte dei casi, questi giochi si basano su un approccio metodologico e teorico molto robusto, il role-play o gioco di ruolo, traslato in ambiente digitale. Questa tecnica, già ampiamente utilizzata nei percorsi formativi tradizionali, è considerata infatti tra i migliori metodi per l’acquisizione e il miglioramento delle competenze trasversali, proprio per le caratteristiche relazionali tipiche delle soft skill. Il loro potenziale per l’apprendimento emerge sia nella loro versione originale, in presenza, che in quella, più recente, in ambienti digitali, con

l'ausilio delle nuove tecnologie. Le caratteristiche e i vantaggi offerti da queste nuove tecniche sono descritte nei paragrafi successivi.

1.3.1 I giochi di ruolo per la formazione delle soft skill

Originariamente il gioco di ruolo (GdR) o role play game (RPG) nasce come un gioco da tavolo, caratterizzato dalla creazione da parte di un gruppo di giocatori di uno spazio immaginario, dove si sviluppano storie fittizie mediante lo scambio dialettico tra diversi personaggi interpretati dai singoli partecipanti. In ogni sessione di gioco, il master (o narratore) definisce l'ambientazione, le caratteristiche dei personaggi e le regole formali. Il role play non prevede il raggiungimento di un obiettivo specifico e di conseguenza non prevede un vincitore, ma si presenta come un gioco di narrazione e cooperazione. Il GdR, così definito, è diventato rapidamente un gioco in presenza e successivamente, grazie agli sviluppi della tecnologia, ha trovato una maggiore espressione in rete prima con i GdR on line, e poi con i Videogiochi di ruolo online (tra cui i MMORPG, Massively Multiplayer Online Role-Playing Game), che vedono la partecipazione di un numero consistente di giocatori di ogni parte del mondo.

Per le sue caratteristiche, questa tecnica è stata facilmente trasferita e utilizzata nella formazione in presenza, poiché fornisce molteplici stimoli all'apprendimento attraverso l'imitazione, l'azione, l'osservazione del comportamento degli altri e i commenti ricevuti sul proprio, oltre che un'analisi dell'intero processo da parte un tutor o master. Ciò consente una successiva analisi dei vissuti, delle dinamiche interpersonali, delle modalità di esercizio di specifici ruoli, e più in generale dei processi di comunicazione agiti nel contesto rappresentato. Sul piano dell'apprendimento, il gioco di ruolo permette ai partecipanti di amplificare la loro flessibilità mentale, l'immaginazione, allargare le prospettive personali e creare più punti di vista, estendere pensieri e sentimenti oltre l'orizzonte di interpretazione personale. Boggs (2007) ritiene il GdR in grado di alimentare le competenze e le capacità cognitive ed emotive dei soggetti, che grazie al gioco si aprono a nuove soluzioni e opportunità identificando e affrontando un nuovo livello di realtà e interpretazione dell'ambiente circostante, sia in relazione a loro stessi che nell'interazione con gli altri.

Le prime forme di GdR come strumento con finalità diverse dal gioco, sono state introdotte da Moreno nel 1934, che propose lo psicodramma e il sociodramma come tecniche terapeutiche basate sulla messa in scena di storie immaginarie guidate da uno psicologo e finalizzate rispettivamente alla risoluzione di problemi personali, nel primo caso, e di problemi collettivi, nel secondo (1977). Il successo di queste tecniche è mostrato anche dal loro utilizzo in altri approcci, a seguito di integrazioni e adattamenti, come ad esempio nelle tecniche Gestaltiste o nelle terapie di gruppo psicoanalitiche. Ben presto però lo stesso Moreno si rese conto delle potenzialità di questa metodologia affermando che queste tecniche dovevano essere messe a disposizione di un pubblico più vasto per scopi diversi dalla clinica. È in questa atmosfera che nascono le prime applicazioni del metodo psicodrammatico classico per fini didattici, per la crescita personale o per migliorare le relazioni interpersonali, ossia in tutti quei campi, diversi dalla psicoterapia, che richiedono un minimo di esplorazione psicologica per la risoluzione di un problema. Più nello specifico, il gioco di ruolo è stato applicato per la formazione dei più svariati argomenti: con i bambini per apprendere una specifica materia (Hollander e Hollander 1978), per lo sviluppo di valori sociali (Shaftel e Shaftel 1967), per l'acquisizione di abilità interpersonali (Schonke, 1975) per potenziare la creatività in classe (Torrance e Myers, 1970) o come vero e proprio metodo di insegnamento nelle scuole (Shaftel e Shaftel 1982). Di più recente introduzione è invece il gioco di ruolo nell'ambito professionale e organizzativo. L'adozione di questa tecnica in questi settori va di pari passo all'evidenza dell'importanza delle soft skill per il potenziamento delle dimensioni interpersonali e intrapersonali degli individui (vedi paragrafo 1.2) e del ruolo che esse giocano nel buon funzionamento di un'azienda, sia in termini totalitari che di singoli gruppi o persone. Blatner (2006), per esempio, sostiene che il gioco di ruolo è il modo migliore per sviluppare competenze trasversali, quali l'iniziativa, la comunicazione, il problem-solving, la consapevolezza di sé e la cooperazione. Questa considerazione è sostenuta anche da Bollens e Marshall (1973), i quali mostrano come il GdR trovano particolare applicazione nei contesti relazionali, in genere in settori quali la consulenza e soft skill training. La formazione può essere riferita ad un tema specifico, come la leadership, la negoziazione, il processo decisionale, o combinazioni di queste, che vengono trasferite in attività di role-play

esperienziale, il cui obiettivo è quello di completare l'acquisizione passiva di conoscenze, aiutando gli studenti a padroneggiare gli aspetti pratici e le dimensioni relazionali. Di fatto, quello che rende questa tecnica particolarmente adatta a contesti professionali, è che si basa sul concetto, già descritto precedentemente, del learning by doing, sperimentare cioè l'apprendimento in contesti pratici in maniera attiva. Gli utenti possono mettere in pratica una serie di comportamenti in ambienti non giudicanti senza la possibilità di sbagliare, e crescere grazie a momenti di riflessione collettiva o di auto-riflessione. Questo processo aiuta a sviluppare una buona flessibilità mentale e un adattamento creativo alle diverse situazioni, poiché sviluppa un tipo di competenza che trascende la conoscenza procedurale e, allo stesso tempo promuove la generalizzazione dei comportamenti e delle nozioni a situazioni diverse da quelle che sono state oggetto di gioco (Perkins e Salomon, 1992).

Oltre alla formazione, il role-play viene utilizzato anche per la selezione del personale in combinazione con una varietà di altri strumenti quali le interviste, test "carta e matita" e casi aziendali, all'interno di quello che viene normalmente definito **Assessment Center** (vedi paragrafo 1.1.3). In questo caso il GdR consente di *ampliare le conoscenze dei soggetti, indagando il comportamento di un individuo in una serie di situazioni ipotizzate, come le modalità di risoluzione di una situazione problematica o le capacità personali per superare lo stress e l'ansia sul posto di lavoro*. L'uso di giochi di ruolo nei processi di valutazione aiuta a identificare e valutare le prestazioni di un dipendente sia individualmente che collettivamente, così come la sua capacità di assumere posizioni e responsabilità superiori, fornendo informazioni utili sulle competenze chiave possedute dal soggetto, come la decision making, la leadership, la capacità di pianificazione e il pensiero analitico, considerate fondamentali per una buona performance lavorativa.

Tuttavia, indipendentemente dagli scopi che guidano il gioco di ruolo, in tutti gli ambiti assume fondamentale importanza il processo di debriefing, ossia una sessione di post-gioco che consente ai giocatori, osservatori, e formatori di condividere i sentimenti esperiti, per esaminare i comportamenti agiti ed elaborare le esperienze individuali e di gruppo (Kozma et al., 1978). Come Thatcher (1990), Petranek e colleghi (1992) hanno sottolineato come il valore del processo di debriefing sia strettamente connesso alla possibilità di discussione tra i diversi attori coinvolti,

formatori, tirocinanti o studenti, finalizzata all'identificazione e alla definizione degli obiettivi personali e strategici e che consente così di riflettere sul proprio processo di apprendimento.

In sintesi, quindi, il role-play è ampiamente riconosciuto come una potente tecnica per migliorare la formazione tradizionale delle competenze trasversali, poiché aumenta l'esperienza di apprendimento dei partecipanti, facilita la conoscenza e l'attuazione di specifiche abilità e competenze, sia sul piano individuale che sociale (McGill & Beaty, 1995; Shaw, et al, 1980 Turner, 1992;. Van Ments, 1999). Diversamente dai metodi più tradizionali, i giochi di ruolo sembrano stimolare la motivazione e l'interesse degli studenti, determinando così un impatto positivo sul cambiamento degli atteggiamenti (Pierfy 1977; Bredemeier & Greenblat, 1981; Van Sickel, 1986), e forniscono un più efficace apprendimento comportamentale, cognitivo ed affettivo, grazie ad una maggiore comprensione personale dei problemi sociali e degli eventi circostanti (Bredemeier & Greenblat, 1981; Foster et al, 1980;. Hankinson, 1987). Per questi motivi, la tecnica del gioco di ruolo è stata adattata a diversi contesti e applicata in molti ambiti disciplinari, come la psicologia, l'organizzazione, la sociologia e la pedagogia, e ampiamente usata per la formazioni di specifiche professioni come infermieri, insegnanti, poliziotti, medici e così via. Questa tendenza è diventata, negli utili anni, ancora più virale grazie al crescente sviluppo delle nuove tecnologie, che hanno permesso la realizzazione di nuove forme e modalità di formazione e valutazione delle soft skill, mediante l'ausilio di strumenti e tecniche proprie del mondo dei video giochi.

1.3.2 I giochi di ruolo vs serious game

In ambito formativo, l'innovazione tecnologica ha permesso di realizzare diverse esperienze concrete di giochi educativi e strumenti di formazione applicata ad una varietà di competenze trasversali, come la negoziazione, il processo decisionale, la leadership, e di problem solving, permettendo la realizzazione di quelli che sono definiti EMORPG (Educational Massive Online Role Playing Game), tecnologia che consente a più individui, rappresentati da avatar, di partecipare simultaneamente, in tempo reale, a sessioni di formazione on-line nell'ambiente virtuale.

In letteratura si evidenzia un interesse crescente sull'importanza di questo tipo di giochi in quanto sono in grado di realizzare esperienze impegnative, motivanti e interessanti mantenendo alta l'attenzione degli utenti che, come in qualsiasi contesto formativo, hanno sempre bisogno di essere sostenuti attraverso un sistema di feedback, di riflessione e di transfert in situazioni di vita reale (Squire, 2003; Egenfeldt-Nielsen, 2005; Michael e Chen, 2005; de Freitas e Oliver 2006). Si può affermare che le considerazioni sull'utilizzo dei giochi per scopi formativi è radicalmente cambiata negli ultimi anni, soprattutto rispetto alla percezione sull'efficacia e la funzionalità dei giochi applicati a questo ambito (Egenfeldt-Nielsen 2007). Sempre più, i ricercatori, psicologi, insegnanti, educatori e sviluppatori di giochi hanno riconosciuto l'importanza di lavorare insieme per disegnare un sistema di principi che aiutino la progettazione di giochi come strumenti di apprendimento efficaci. In questo campo, giochi di simulazione di ruolo possono essere concepite come strumenti che, se integrati opportunamente in programmi di formazione, possono aumentare efficacemente l'apprendimento.

In letteratura non esiste una definizione univoca, ma vengono utilizzati termini diversi come serious games, simulazioni, mondi virtuali, giochi di ruoli digitali. Tuttavia, questi sistemi si possono considerare come punti su una linea continua, in quanto si differenziano fondamentalmente per la struttura del gioco e la specificità degli obiettivi di apprendimento (Aldrich, 2009; Carr et al 2010). Inoltre, essi possono essere tutti considerati delle attività di apprendimento esperienziale o metodi attivi (Dewey, 1938, 1966) utilizzati per lo sviluppo di diverse competenze (conoscenza procedurale, competenze tecniche, professionali, personali e sociali), con finalità distinte e in contesti diversificati (ad es. sanitario, militare, educativo, formativo e professionale).

Abbiamo già visto nel paragrafo precedente (1.3) come si sono differenziati i giochi on line dalla loro nascita ad oggi. Si distinguono i giochi online multigiocatore di massa, MMORPG (vedi paragrafo 1.3.1) basati sulla tecnologia role-play con finalità prettamente ludiche, da quelli che sebbene nati per il divertimento vengono poi utilizzati anche per motivi di apprendimento (Serious Gaming), fino a quelli con scopi esclusivamente educativi, i cosiddetti Serious Game. Esemplicando, rientrano nella prima categoria giochi come Second Life, ossia un mondo virtuale immersivo

per un numero enorme di giocatori, nato per scopi sociali e di puro intrattenimento. Tuttavia lo stesso Second life o suoi adattamenti, come SciLands e VirtualPREX2, ossia ambienti all'interno del gioco primario pensati per l'approfondimento di specifiche tematiche quali scienza e tecnologia nel primo caso e insegnamento nel secondo, sono stati utilizzati con fini educativi (Kirriemuir, 2008; De Freitas 2008, Tojde, 2011, Gregory & Masters, 2012) rientrando così nella seconda categoria. Fanno parte di questa categoria anche i giochi multipiattaforma con forme di apprendimento autogestite come ad esempio Croquet, un sistema che permette agli utenti di creare mondi virtuali per l'insegnamento di diverse discipline, come le lingue, la storia, l'architettura, basandosi esclusivamente sulla comunicazione e la condivisione delle risorse da parte di soggetti coinvolti; e Open Wonderland dove gli utenti interagiscono attraverso avatar che comunicano tra loro con sistemi audio o testuali (le chat), particolarmente utilizzato per migliorare le competenze di team working in ambito lavorativo. Fanno parte invece dell'ultima tipologia tutti quei giochi progettati esclusivamente per la formazione, caratterizzati da una serie di scenari di apprendimento on-line montati su piattaforme e-learning, i cui contenuti sono progettati, gestiti e valutati da tutor, formatori o psicologi. Come abbiamo già detto, si tratta di sistemi EMORPG (vedi sopra), ossia giochi esperienziali ispirati ai principi di ruolo, realizzati in un mondo virtuale dove gli utenti interagiscono attraverso rappresentazioni grafiche dette avatar e sotto la supervisione e la guida di un regista. Il regista (che a seconda del contesto può essere uno psicologo, un insegnante, un formatore, un educatore, un consulente o un terapeuta) può svolgere ruoli diversi, tra cui scrivere uno storyboard, assegnare i ruoli ai giocatori, guidare l'interazione, commentare la performance del gruppo, registrare e analizzare le parti significative della sessione di gioco per fornire dei feedback adeguati (debriefing). In questo modo, tutti i sistemi EMORPG forniscono una struttura, un disegno della esperienza, processi di feedback, per la riflessione e la presa d'atto delle proprie azioni, assicurando così l'acquisizione e il trasferimento delle competenze oggetto di studio. Come esempio di giochi di questo tipo, possiamo citare E-Adventure, una piattaforma spagnola per l'insegnamento della storia e della lingua inglese, E-Circus pensato per la creazione di percorsi personalizzati finalizzati all'apprendimento sociale ed emotivo e disponibile sia in modalità single che multiplayer, Fear-Not

specifico per combattere il bullismo e aumentare l'empatia interculturale nelle scuole, e InfiniTeams pensato invece per contesti professionali e finalizzato all'acquisizione di competenze trasversali utili sul lavoro, come il team-building, la leadership, la fiducia, la negoziazione e il problem solving. Un'intensa attività in questo senso è stata svolta anche dal laboratorio NAC, Natural and Artificial Cognition Lab (www.nac.unina.it), dell'università di Napoli Federico II che negli ultimi 10 anni, grazie a diversi progetti finanziati dalla Comunità Europea, ha realizzato una serie di serious games, che coniugano tecnologia, psicologia e pedagogia, per l'acquisizione di specifiche soft skill in contesti definiti. Tra questi, di particolare interesse:

- 1) Learn to Lead (www.l2l.unina.it): serious game per insegnare/apprendere le teorie della leadership in ambito aziendale.
- 2) Decide-it (www.decide-it.eu): serious game per insegnare/apprendere la presa di decisione in condizioni di emergenza in ambito aziendale.
- 3) Dread-ed (www.nac.unina.it/nac/progetti): serious game per la gestione delle emergenze e catastrofi naturali, utilizzati per istituzioni, scuole e aziende.
- 4) S-Cube (www.s-cube.eu): piattaforma per la progettazione e l'erogazione di giochi di ruolo educativi sulle soft skill aziendali.
- 5) ENACT (www.enactskills.eu): serious game per la formazione e valutazione delle competenze di negoziazione.

Come abbiamo visto, i serious game hanno un riscontro notevole, sia in termini di progettazione e sviluppo, che di fruizione. Tuttavia, un'attenta analisi della letteratura, mostra che insieme ai numerosi vantaggi e potenzialità (vedi paragrafo 1.3), questi sistemi mostrano alcune criticità legate soprattutto alla possibilità di valutare i risultati dell'apprendimento offerto e più specificatamente all'efficacia e all'efficienza delle metodologie utilizzate. È in questo panorama che hanno preso avvio gli studi sull'assessment applicato ai SGS, dapprima come forma di valutazione di queste nuove tecnologie di formazione, e in tempi ancor più recenti, per la realizzazione di giochi progettati specificatamente per la valutazione di competenze specifiche.

1.3.3 I Serious Game come strumenti di Assessment

In questo nuovo panorama, la prima sfida degli addetti ai lavori era quella di trovare un set di metodologie e di strumenti in grado di valutare l'apprendimento ottenuto attraverso i SGS, tenendo conto sia della dimensione ludica che dell'impatto educativo. La necessità di valutare i risultati di apprendimento ottenuti "giocando", risulta infatti essenziale per i) progettare SGS più efficaci ed efficienti; ii) identificare quali SGS sono i più adatti per un dato obiettivo o dominio, iii) fornire feedback o personalizzare il gioco in maniera adeguata alle performance del giocatore, al fine di garantire la motivazione e l'apprendimento stesso.

Ma cosa si intende per efficacia di un Serious Game? Secondo Michael and Chen (2005), i SGS, come ogni altro strumento di educazione, devono essere in grado di dimostrare che si è verificato l'apprendimento necessario, ossia che gli obiettivi di apprendimento sono stati soddisfatti. Tuttavia, l'apprendimento è un costrutto complesso, già di per se difficile da misurare, e determinare se una simulazione o SGS sia efficace al raggiungimento degli obiettivi di apprendimento previsti è un processo complesso, lungo, costoso e difficile. Parte di questa difficoltà deriva dalla natura dei videogiochi che rende complessa la raccolta di dati.

Come afferma G. Bente (2009), nel valutare i SGS è difficile sia dimostrare che gli studenti stanno imparando quello che dovrebbero imparare e, soprattutto, sapere che quello che si sta misurando è esattamente quello che si pensa che si stia misurando. Queste considerazioni hanno creato un vero e proprio asse di interesse, relativo alla valutazione, sia sul piano teorico che metodologico.

Tradizionalmente la valutazione può essere di due tipi: Sommativa, al termine del processo di apprendimento; o Formativa, effettuata durante l'apprendimento. Nell'ambito dei Serious Game, Michael e Chen, nel 2005, identificano invece tre principali tipi di valutazione:

1. Completion ass (equivalente alla valutazione sommativa): misura se il risultato finale è stato ottenuto.
2. In-process ass (equivalente alla valutazione formativa): misura come, quando e perché il giocatore ha fatto delle scelte.

3. Teacher ass.: osservazioni e giudizi dell'insegnante in azione, difficilmente catturabili da una tecnologia.

La valutazione formativa, o in-process, è quella più auspicabile per i SGS, ma la più difficile da realizzare, rispetto a quella sommativa, che è per tanto la più utilizzata. Normalmente infatti nei contesti formativi con i SGS vengono utilizzati: 1. Pre-post test, per misurare l'acquisizione della competenza e/o verificare l'effetto di un nuovo metodo di insegnamento, il cui limite principale è dato dalla difficoltà di controllare gli effetti delle variabili esterne. 2. Self-report, questionario/test finale di auto-percezione del giocatore, il cui principale limite è dato dal fatto che si basa sull'opinione dell'utente e non su quanto realmente accaduto nel gioco. Entrambe sono misure dirette, tipiche della valutazione sommativa.

Esistono, tuttavia, altre misure che possono fornire indicazioni sull'efficacia dei SGS, sono le cosiddette misure indirette, quali ad esempio, il livello di impegno, di motivazione, ecc.. che indicano il grado di interesse creato nell'utente; o le misure fisiologiche, quali ad esempio l'attività muscolare facciale, misure cardiovascolari, galvaniche e cerebrali, che forniscono indicazioni fisiologiche sugli effetti del gioco sull'utente ma che sono naturalmente molto difficili da impiegare. Infine, le misure derivate dall'utilizzo di tecnologie audio-visive, come videoregistrazioni accompagnate da debriefing con l'insegnante, che offrono la possibilità di raccogliere una vasta gamma di informazioni, sia esplicite che implicite.

Quello a cui si è assistito negli ultimi 10 anni, è un generale aumento di interesse in questo settore che si è tradotto nella realizzazione di un numero sempre maggiore di software per l'assessment, finalizzati a misurare in maniera automatica e in tempo reale le prestazioni degli utenti. Questi software si differenziano per tipologia e per obiettivo. Tra i più diffusi possiamo identificare: 1. Assessment Tool for Teaching and learning (e-asTTle), specifico per compiti di lettura, scrittura, e matematica; 2. Questionmark Perception, piattaforma per la creazione di test, questionari, quiz su specifici argomenti; 3. International Knowledge Measurement, servizio online per la profilazione di candidati in contesti lavorativi.

Nonostante le differenze, bisogna però sottolineare che questi strumenti sono tutti accumulati dal carattere *non* innovativo della misurazione. In sostanza infatti si tratta

di strumenti che in maniera più sofisticata, rapida ed economica, non fanno altro che sostituire i vecchi metodi carta e matita, quali questionari, checklist, test e quiz, ma che non ne modificano l'approccio. Come i metodi tradizionali, anche questi software presentano infatti limiti noti, come ad esempio interrompono il gioco, demotivano gli studenti, annoiano, e così via, e non offrono nulla di veramente nuovo, nonostante la formazione fornita con i serious game sia del tutto diversa rispetto a quella tradizionale.

Per questo motivo, i lavori più recenti si stanno orientando su metodi di valutazione implementati nel gioco stesso, che forniscono misure oggettive delle performance dei giocatori, tenendo conto delle diverse fasi della formazione e del gioco, in grado di permettere una personalizzazione delle attività svolte in base alla performance misurata in progress, e che forniscono feedback sulla prestazione in tempo reale. I primi esperimenti in questo campo sono rappresentati da CancerSpace (2010), sviluppato per il settore sanitario e PIXELearning Enterprise game (2010), per l'ambiente business. Si tratta di lavori pionieristici, lontani da quello che effettivamente si teorizza sia necessario per attuare questa nuova visione di valutazione. Questo perché l'integrazione della valutazione all'interno del gioco è una questione tutt'altro che semplice, che richiede studi approfonditi sia sul piano teorico/metodologico che tecnologico, al fine di rendere l'esperienza di apprendimento definibile, quantificabile e misurabile (K. Corti della PixelLearning).

Uno di questi tentativi è rappresentato proprio da **ENACT**, l'ultimo nato nella lunga lista di prodotti realizzati dal laboratorio NAC, realizzato in collaborazione con diversi partner stranieri, sia pubblici che privati, rientra proprio in questa tipologia, in quanto **propone**, oltre a formazione, **un sistema di valutazione delle competenze insito nel gioco stesso**, elemento che come abbiamo visto risulta assolutamente innovativo. Ed è proprio **su questo strumento che ci incentra il presente lavoro, descritto nel dettaglio nel capitolo 3, la Ricerca.**

I capitoli che seguono, forniscono una descrizione dettagliata del sistema, partendo dall'approccio teorico di riferimento, il modello dei 5 fattori di gestione dei conflitti di Rahim, fino alla descrizione dei risultati.

2 Il punto di partenza: il ROCI e La teoria di Rahim

Dalla breve overview presentata nel capitolo precedente, risulta evidente il forte interesse da parte dei ricercatori verso la valutazione delle soft skill e il tentativo di sviluppare strumenti tecnologici innovati tesi a sostituire i “tradizionali metodi carta e matita”. Il presente studio rientra in questo quadro, in quanto indagata la possibilità di integrare le teorie e le metodologie della valutazione psicologica con le tecnologie dei Serious games, attraverso la progettazione, la realizzazione e la validazione di un software di formazione e valutazione specifico per le capacità di negoziazione: ENACT, Enhancing Negotiation skills through on-line Assessment of Competencies and interactive mobile Training. La descrizione dettagliata dello strumento è riportata nel capitolo successivo, mentre in questo capitolo sono definiti i concetti e le teorie utilizzate per la sua realizzazione e le motivazioni sottese alle scelte metodologiche che sono state fatte.

Innanzitutto, la soft skill su cui si incentra questo lavoro è la **negoziazione**, in quanto considerata una competenza sociale intra-interpersonale fondamentale per l'apprendimento e lo sviluppo personale, dato il suo forte impatto sulla vita personale e professionale degli individui. La capacità di negoziare è riconosciuta come un elemento chiave nella formazione imprenditoriale e, più in generale, delle competenze aziendali (Frank, 2007), risulta fondamentale al fine di ottenere una migliore inclusione sociale ed economica (Turok & Taylor, 2006), oltre che per migliorare la propria prestazione lavorativa (Baron & Markman, 2003). Per queste sue caratteristiche, la stessa Commissione Europea ha inserito la negoziazione tra le competenze chiave per l'apprendimento permanente in quanto capace di migliorare sia le "competenze sociali e civili" sia la capacità imprenditoriale" degli individui (vedi paragrafo 1.2.2).

Anche se è estremamente difficile individuare una definizione unica e universalmente accettata di negoziazione, dati i molteplici usi fatti negli svariati campi di applicazione (tra cui economia, diritto, relazioni internazionali, psicologia, sociologia e gestione dei conflitti), in questo studio si fa riferimento all'**approccio integrativo** (Walton e McKersie, 1965; Fisher e Ury, 1981; Zartman e Berman,

1982) che definisce la negoziazione come un processo cooperativo e non competitivo-distributivo. Quest'ultimo definisce la negoziazione come una situazione di conflitto dove le diverse parti competono per ottenere il miglior risultato, ossia ciascun attore coinvolto mira non solo ad ottenere il massimo vantaggio per se stesso, ma anche a lasciare all'altro il meno possibile. *L'approccio integrativo invece ha come obiettivo la ricerca di una soluzione integrata, basata sulla possibilità di "fare concessioni" al fine di raggiungere un accordo, che sia il più soddisfacente per tutti i soggetti coinvolti.* In questo quadro, pertanto, **la negoziazione è vista come processo di feedback costruttivo ed efficace, volto alla ricerca di una soluzione comune di fronte alle diverse esigenze o idee personali.** Un processo in cui l'altra parte non è più intesa come un avversario, ma come un partner, grazie alla capacità di “mettersi nei panni di”, di imparare ad adottare posizioni altrui. Ma non esiste una negoziazione efficace senza una comunicazione efficace ed un'efficace consapevolezza di sé. Fisher e Ury (1981) sostengono che, in una trattativa, una percezione chiara di sé e dell'altro può essere un elemento fondamentale per entrambe le parti. Questo perché una comunicazione efficace può cambiare gli atteggiamenti, prevenire o superare le incomprensioni e contribuire a migliorare le relazioni. È noto infatti che le barriere comunicative ostacolano il processo di negoziazione. L'approccio integrativo si concentra sulla costruzione di relazioni di fiducia reciproca tra le parti e queste vengono realizzate attraverso un atteggiamento di comunicazione efficace.

2.1 Rahim e la gestione dei conflitti

Nell'ambito dell'approccio integrato, il **modello teorico di riferimento** che si è scelto di utilizzare per la definizione e misurazione della capacità di negoziare è “**Il modello dei cinque fattori di gestione dei conflitti interpersonali**” di **Rahim & Bonoma (1979)**. Questa scelta si basa sull'approccio proprio degli autori, che non ritengono il conflitto un concetto negativo di per sé, ma sostengono che se gestito correttamente può rappresentare un'ottima opportunità per crescere, imparare e lavorare ancora più efficacemente. Pertanto, la capacità di gestire i conflitti è strettamente connessa alle capacità di negoziazione proprie dell'individuo. Inoltre, il concetto di conflitto organizzativo proposto da Rahim costituisce il più rilevante punto di riferimento nell'ambito della letteratura di psicologia organizzativa. Questo

soprattutto in ragione degli strumenti elaborati dall'autore: il R.O.C.I. I e II (Rahim Organizational Conflict Inventory I e II), i test sulla gestione del conflitto tra i più utilizzati al mondo, il primo pensato per misurare le dimensioni del conflitto all'interno delle organizzazioni, il secondo per misurare i modi di affrontare il conflitto interpersonale. Secondo Rahim (1995), esistono **tre diverse tipologie di conflitto**, ciascuna innescata da fattori diversi:

- 1) Intrapersonale: si tratta di un conflitto individuale che nasce ogni qualvolta non c'è corrispondenza tra le competenze attese, ovvero i compiti assegnati dall'organizzazione all'individuo, con le competenze individuali, intese come insieme di interessi, valori e capacità personali. Le cause di questo conflitto sono prevalentemente di tipo strutturale (Majer, 1995) dovute ad un'incompatibilità tra la persona e il compito; oppure tra i bisogni della persona e gli obiettivi organizzativi; o anche a causa di richieste eccessive da parte dell'organizzazione rispetto alle reali capacità della persona. Un elevato conflitto intrapersonale determina spesso un aumento dell'assenteismo così come di malattie e disturbi psicosomatici.
- 2) Intra-gruppo: in questo caso il conflitto coinvolge i membri di un gruppo e si svolge all'interno di esso. I fattori scatenanti possono essere determinati i) da uno stile di leadership autoritario, ii) da una struttura del compito troppo complessa, iii) dalle grandi dimensioni del gruppi di lavoro, iv) dalla diversità di atteggiamenti, valori, interessi, stili interpersonali all'interno del gruppo; e v) da risultati negativi in ambienti ad elevata competitività interna.
- 3) Inter-gruppi: il conflitto emerge tra gruppi di lavoro diversi come ad esempio tra squadre, tra reparti, tra funzioni di una stessa organizzazione. Questo genere di conflitto si evidenzia soprattutto in presenza di una forte differenziazione tra sottosistemi di un'organizzazione; o quando si crea un'elevata interdipendenza tra reparti; o più semplicemente quando le risorse sono troppo limitate.

Nell'ambito della presente ricerca sono stati presi in considerazioni **solo i conflitti interpersonali**, in quanto rappresentativi del tipo di conflitti esperiti dagli individui in tutti i contesti di vita.

Per la risoluzione di un conflitto, si possono distinguere, secondo l'autore, cinque diversi stili di gestione, che derivano dalla combinazione di due dimensioni fondamentali (intenzioni) su cui si inserisce la negoziazione: la preoccupazione per sé, ossia il grado (alto o basso) di interesse a soddisfare le proprie esigenze, e la preoccupazione per gli altri, ossia il grado (alto o basso) di interesse a soddisfare le esigenze degli altri. Si può dire che queste dimensioni ritraggono gli orientamenti motivazionali di un dato individuo durante un conflitto (Rubin & Brown, 1975), e sono supportate da una serie di studi, tra cui quelli di Ruble e Thomas (1976) e Van De Vliert e Kabanoff (1990). La combinazione delle due dimensioni si traduce in 5 specifici stili di gestione del conflitto interpersonale, come mostrato in Figura 4, che secondo l'autore, ogni individuo dovrebbe conoscere ed essere in grado di padroneggiare in base alla situazione conflittuale che si trova ad affrontare.

		CONCERN FOR SELF	
		High	Low
CONCERN FOR OTHER	High	INTEGRATING apertura, scambio di informazioni, valorizzazione delle differenze per raggiungere soluzioni accettabili per entrambe le parti	OBLIGING sdrammatizzare le differenze, enfatizzando gli aspetti comuni per soddisfare le preoccupazioni degli altri
	Low	DOMINATING Orientamento win-lose, comportamento teso a soddisfare la propria posizione	AVOIDING ritiro, scarica barile, aggiramento delle situazioni

COMPROMISING
 dare-e-prendere, in cui entrambe le parti rinunciano a qualcosa per rendere la decisione accettabile

Figura 4: Il modello di Rahim

Integrating (IN): questo stile è caratterizzato da un alto punteggio su entrambe le dimensioni, che si esprime in una volontà di collaborazione tra le parti, con scambio di informazioni, apertura ed analisi delle differenze. Questa modalità è particolarmente appropriata nelle situazioni in cui i risultati da raggiungere sono complessi, ed è necessaria una collaborazione per raggiungere un risultato, o quando c'è tempo a disposizione per risolvere il problema, oppure quando è necessario utilizzare le risorse di entrambe le parti. E' inappropriata nelle situazioni in cui il problema è semplice e quando è necessaria una risposta immediata.

Obliging (OB): caratterizzato da una maggiore preoccupazione per gli altri, con questo stile il soggetto minimizza le differenze e cerca di enfatizzare la comunanza

di interessi. E' appropriato nelle situazioni in cui si pensa di essere in torto, oppure quando il risultato è più importante per l'altro, o ancora, quando strategicamente, si concede qualcosa ora per ottenere qualcos'altro in futuro. Inoltre dovrebbe essere adottato quando ci si trova in una posizione di debolezza o quando si privilegia il mantenimento dei buoni rapporti. E' inappropriato nelle situazioni in cui il risultato è importante per se stessi o si ritiene di essere nel giusto.

Dominating (DO): chiaramente incentrato sulla preoccupazione per sé, è caratterizzato da un perseguimento ostinato dei propri obiettivi ignorando quelli degli altri. E' appropriato in quelle situazioni in cui c'è bisogno di decisioni rapide o quando il soggetto ha una maggiore competenza sull'argomento. E' inappropriato quando il risultato è complesso, per ottenere qualcosa di scarsa importanza, o in presenza di una controparte potente e competente.

Avoiding (AV): caratterizzato da un basso punteggio su entrambe le situazioni, con questo stile si evita il confronto, la disputa, accantonando la questione, con il risultato di un vero e proprio ritiro dal conflitto. E' appropriato nelle situazioni in cui il risultato è di poco conto, quando è necessario prendere del tempo per calmarsi, quando affrontare l'altro è più disfunzionale dell'utilità del risultato. Invece risulta inappropriato nel caso in cui il risultato è importante, o quando è necessario assumersi la responsabilità di una decisione, o quando il problema necessita di attenzione immediata.

Copromising (CO): si tratta di uno stile intermedio, caratterizzato da una moderata preoccupazione per sé e per gli altri, e si manifesta con una rinuncia da entrambe le parti pur di raggiungere una decisione reciprocamente accettabile. E' appropriato quando entrambe le parti sono potenti alla stessa maniera o quando è impossibile raggiungere un consenso unanime o anche quando gli altri stili, quali DO e IN, non hanno avuto successo, o se è necessaria trovare una soluzione temporanea ad un problema complesso. E' inappropriato invece quando una parte è molto più potente.

Le cinque modalità della gestione dei conflitti teorizzati da Rahim sono valutati attraverso una scala realizzata dallo stesso autore: il ROCI II.

2.2 II ROCI II

Il **ROCI II** è il test, sviluppato da Rahim, che sostiene le sue teorie da un punto di vista scientifico. Questo test è il più diffuso e rigorosamente testato strumento per la valutazione dei conflitti, raccomandato per la rilevazione dell'entità del conflitto e nell'individuazione delle vie da seguire per rendere il conflitto stesso non disfunzionale o se non addirittura funzionale ai risultati che l'individuo persegue. Lo strumento contiene tre diversi moduli A, B, e C per investigare gli stili di gestione dei conflitti rispettivamente con i superiori, subordinati e pari, all'interno di una organizzazione. Nell'ambito della ricerca si farà riferimento alla **FORMA C - conflitti interpersonali tra coetanei**, il cui testo integrale è riportato nell'allegato 1.

Il test si compone di 28 item che misurano le 5 dimensioni indipendenti degli stili di gestione del conflitto, di cui sette formano la sottoscala Integrating, sei quella Obliging, cinque per Dominating, sei per Avoiding, e quattro item per Compromising. Gli item sono stati selezionati in seguito ad una serie ripetuta di analisi degli item e analisi fattoriale. Tutti gli item sono misurati su scala Likert a 5 passi che va da "fortemente in disaccordo" (1) a "molto d'accordo" (5). Un punteggio alto rappresenta un maggiore uso di uno stile di conflitto. La Tabella 4 propone esempi degli item delle diverse sottoscale. Si tratta di un test breve, richiede infatti solo 8 minuti per la sua compilazione, ed svolto in autonomia senza limiti di tempo.

Item n.	Scala	Descrizione
5	INT	<i>Cerco di lavorare con i miei colleghi per trovare una soluzione a un problema che soddisfa le aspettative di tutti.</i>
24	OBL	<i>Cerco di soddisfare le aspettative dei miei coetanei.</i>
8	DOM	<i>Uso la mia influenza per far sì che le mie idee siano accettate</i>
27	AVO	<i>Cerco di evitare spiacevoli scambi con i miei coetanei.</i>
7	OBL	<i>Cerco di trovare un compromesso per risolvere le situazioni di stallo</i>

Tabella 4: Esempi di item delle 5 scale del ROCI

Rahim ha condotto numerosi studi sul ROCI II, dimostrando che le 5 sottoscale, una per ogni stile, hanno un'adeguata affidabilità e validità (vedi paragrafo 1.1.2). Nel 1995 ha condotto un'analisi fattoriale confermatrice del test su ben 5 campioni

diversi coinvolgendo più di 2.000 soggetti (C1=484 studenti con un lavoro a tempo pieno, C2=550 amministratori pubblici; C3=214 amministrativi universitari; C4=250 dipendenti e direttori di banca in Bangladesh; e C5=578 dirigenti e dipendenti). L'obiettivo era di valutare la validità di costrutto delle cinque sottoscale del ROCI-II e il loro fattore di invarianza tra i gruppi. I risultati dimostrano una buona validità convergente e discriminante delle sottoscale di tutti e 5 gli stili di gestione dei conflitti, e un buon risultato anche per lo studio dell'invarianza del modello a 5 fattori sia tra i gruppi (superiori, subalterni e pari), sia tra i livelli organizzativi (alto, medio, basso e non management), che su 4 dei 5 campioni¹¹.

La versione italiana del ROCI II è curata da Majer (1995), ed è corredata da norme tratte da un ampio campione, 1710 soggetti, rappresentativo del contesto organizzativo italiano.

2.3 Perché il modello di Rahim

Diversi autori hanno realizzato studi e proposto teorie sui modelli di gestione dei conflitti interpersonali, tra cui i più citati Blake & Mouton (1964); Follett (1926/1940); Pšeni ka & Rahim (1989); Rahim (1983, 1992); Rahim & Bonoma, 1979; Rahim & Pšeni ka, 1984; Thomas (1976, 1992); ciascuno caratterizzato di una diversa classificazione degli stili.

Mary P. Follett nel 1940, ad esempio, ha teorizzato un modello 3+2, ossia tre modalità principali di affrontare i conflitti, quali domination, compromise, and integration, e due secondari, quali avoidance and suppression. I primi a presentare invece uno schema di classificazione a cinque stili di gestione dei conflitti interpersonali su due dimensioni sono stati Blake e Mouton nel 1964, distinguendo tra forcing, withdrawing, smoothing, compromising, and confrontation. Queste cinque modalità sono il risultato dell'interazione di due dimensioni correlate all'atteggiamento dei manager all'interno di un'organizzazione: concern for production and concern for people. Successivamente Thomas, nel 1976, ha reinterpretato e raffinato la griglia di Blake e Mouton, ponendo alla base della

¹¹ Rahim, A. & Magner, N. (1995). Confirmatory factor analysis of the styles of handling interpersonal conflict: first-order factor model and its invariance across groups. *Journal of Applied Psychology*, 80(1), 122-132.

classificazione le intenzioni dei soggetti coinvolti nel conflitto che possono essere di due tipi: cooperativeness and assertiveness. Rahim, nel 1983, utilizzando una concettualizzazione simile a quella di Blake e Mouton e Thomas, propone il suo modello e relativo test presentando una classificazione a cinque stili, determinata dalla interconnessione di due dimensioni principali, così come descritto nel paragrafo precedente.

Nel 1988, tuttavia, Weider-Hatfield in seguito ad un'analisi dettagliata della letteratura sul tema, ha concluso che "sebbene la letteratura sul conflitto ha storicamente abbracciato il paradigma dei "cinque stili", studi più recenti indicano che gli individui possono scegliere fra tre, e non cinque, distinti stili di conflitto". Allo stesso modo, Hocker e Wilmot nel 1991, in seguito ad una rassegna della letteratura, affermano che "come le tattiche di conflitto, anche gli stili di gestione del conflitto si possono raggruppare in tre tipi principali: (1) avoidance, (2) competitive (distributive) and (3) collaborative (integrative)". Altri autori invece propongono schemi a due o quattro tipi di modalità. La Tabella 5, sintetizza i principali modelli classificandoli in base al numero degli stili di gestione del conflitto.

N. di stili	Stili	Autori
Due stili	Cooperation e competition	Deutsch, 1949, 1990; Tjosvold, 1990
Tre stili	Nonconfrontation, solution-orientation, e control	Putnam & Wilson, 1982
Quattro stili	Yielding, problem solving, inaction, e contending	Pruitt, 1983
Cinque stili	Integrating, obliging, dominating, avoiding, e compromising	Blake & Mouton, 1964; Follett, 1926/1940; Rahim & Bonoma, 1979; Thomas, 1976

Tabella 5: Tassonomia degli stili di gestione del conflitto (Rahim, A. & Magner, N. (1995))

Di fronte a questa grande varietà di modelli presenti in letteratura **la scelta di utilizzare il modello a 5 fattori di Rahim per la realizzazione di ENACT nasce da una serie di considerazioni.** Innanzitutto, **il modello** proposto da Rahim è **particolarmente adatto per la "formazione" di queste competenze.** Questo perché si basa sull'accezione positiva del termine, che ritiene il conflitto una buona opportunità di crescita, di apprendimento e di miglioramento delle competenze

lavorative, determinando sempre una crescita personale. Saper gestire un conflitto in maniera appropriata significa essere consapevoli di se stessi e delle proprie capacità, riuscire a capire le dinamiche di gruppo e essere in grado di definire il giusto momento per agire. *Questa concettualizzazione si sposa appieno con l'obiettivo del progetto, ossia dare agli utenti la possibilità di esercitare le proprie competenze e acquisire una buona conoscenza delle modalità di interazione proprie e quelle degli altri attraverso la sperimentazione "virtuale" proposta da ENACT.*

Inoltre, Rahim pone l'accento sulla **scelta dello stile giusto a seconda delle situazioni** di conflitto da risolvere (2011). Richiama quindi alla necessità di imparare a gestire le diverse modalità e ad essere flessibili in base alle variabili di contesto, quali l'importanza del risultato, il tempo e le risorse a disposizione, le competenze personali e professionali, ecc... *Per questo motivo, risulta particolarmente adatto alla traslazione in un contesto di gioco digitale, come appunto è ENACT, caratterizzato dalla possibilità di inserire numerosi elementi di contesto e farli diventare parte integrante della formazione.*

Infine, sul piano dell'assessment, il **modello di Rahim è sostenuto dal ROCI II** (Rahim & Bonoma 1983, 2001). Dalla sua pubblicazione, il test è stato utilizzato, oltre che dagli autori, anche da molti ricercatori e selezionatori, facendolo diventare di fatto lo strumento psicometrico sulla gestione dei conflitti più utilizzato, testato, standardizzato e normato a livello internazionale. *Ed è proprio per questo suo rigore psicometrico che il modello di Rahim e il ROCI II, sono stati considerati adatti per lo sviluppo di un serious game specifico per l'assessment, quale ENACT.*

3 La ricerca

Il presente studio, di natura multidisciplinare, si muove nell'ambito della valutazione delle soft skill o competenze personali, e in particolare nell'ambito della valutazione delle competenze di negoziazione. Dalla breve rassegna presentata nel capitolo 1 e 2 si evincono due elementi importanti. Innanzitutto che la **capacità di negoziare** è una tra le più importanti soft skill, in quanto è una competenza relazionale, intra e interpersonale che influisce efficacemente, positivamente e direttamente sulla vita sia personale che professionale delle persone. È considerata una competenza centrale nello sviluppo delle capacità imprenditoriali e comunicative, questo perché possedere una buona capacità di negoziazione aumenta le abilità sociali degli individui, che diventano perciò maggiormente flessibili e capaci di gestire i cambiamenti nelle situazioni e negli ambienti. Il secondo elemento si riferisce ai nuovi sviluppi nell'ambito tecnologico e il massiccio uso della rete internet, che hanno determinato la necessità di creare strumenti sempre più attuali per la formazione e la valutazione di questa competenza, che tengano conto degli interessi degli utenti e dei nuovi approcci forniti dalla ricerca, con particolare riferimento all'ambito dei **Serious Game**. Il capitolo descrive pertanto gli elementi principali della ricerca, che si propone come un tentativo, scientificamente valido, di realizzare un SGS specifico per la valutazione delle capacità di negoziazione e di verificare, con tecniche riconosciute e statisticamente rigorose, la bontà dello strumento proposto. Si tratta di un studio, potremmo dire, pionieristico, ed è per questo che la ricerca si conclude con una serie di considerazioni e suggerimenti per il miglioramento dello strumento nel prossimo futuro.

3.1 Obiettivi

La ricerca, come abbiamo detto, cerca di integrare tre diversi temi: l'assessment, le competenze trasversali e i serious game. Operativamente questo si traduce in un duplice scopo.

Il primo obiettivo è proporre ENACT, un serious game nato nell'ambito di un progetto finanziato dalla Comunità Europea, come un innovativo strumento di

valutazione di queste competenze, alternativo ai tradizionali metodi “Carta e Matita”

ENACT è uno serious games in 3D, single-player, per la formazione e la valutazione delle competenze di negoziazione e di comunicazione in scenari realistici, realizzato dal gruppo di ricercatori del NAC, Natural and Artificial Cognition Laboratory dell'Università Federico II di Napoli, in collaborazione con altri 5 partner europei. Con questo nuovo strumento, si è indagata la possibilità di integrare le teorie e le metodologie della valutazione psicologica (psicodiagnostica) con i Serious games, ossia giochi nati e/o utilizzati per scopi educativi e formativi. Nella prima parte della ricerca, saranno descritte tutte le fasi di progettazione e sviluppo del nuovo strumento di assessment, partendo dall'implementazione dei costrutti teorici da cui prende origine, il modello dei 5 fattori di Rahim (vedi capitolo 2), fino alla verifica della sua validità e attendibilità. L'obiettivo è verificare quanto il serious game può rappresentare una forma parallela del ROCI II. Questa specifica parte del lavoro è stata svolta con il NAC, nel biennio 2014-2015, sotto il coordinamento dell'Università di Plymouth.

Il secondo obiettivo è verificare la bontà del nuovo strumento, attraverso un'analisi critica dei risultati del progetto, finalizzata all'identificazione di un metodo statisticamente valido per il confronto dei due test, basata sull'utilizzo di tecniche di statistica esplorativa dei dati per così individuare gli elementi critici necessari di ulteriori studi e approfondimenti.

L'analisi effettuata alla fine del progetto, infatti, se da una parte ha mostrato alcuni elementi positivi, dall'altra ha fatto sorgere molte domande di natura statistica, legate soprattutto ai metodi di analisi utilizzati vista la diversa natura dei due test, carta e matita vs serious game, e le differenze dei dati, sia in termini di numerosità che di tipologia. L'analisi critica effettuata si propone pertanto di utilizzare strumenti della teoria classica dei test per trovare un metodo valido sia per rappresentare la ricchezza e varietà dell'informazione offerta da ENACT, che per confrontare i due strumenti. Il risultato finale è rappresentato da una serie di considerazioni e/o suggerimenti volti a migliorare questo nuovo test di valutazione, che sebbene ancora in fase prototipale e con alcune criticità, rappresenta comunque un valido e innovativo strumento di

assessment soprattutto in riferimento ai cambiamenti sociali e culturali a cui si sta assistendo negli ultimi decenni. Questa seconda parte del lavoro rappresenta il contributo personale realizzato nell'ultimo anno di dottorato.

3.2 ENACT: un nuovo strumento di valutazione psicologica

Come già sottolineato, una delle caratteristiche che rende le Soft Skill un tema particolarmente interessante è che esse possono essere arricchite nel corso della vita, grazie alle esperienze personali e professionali, e quindi di fatto possono essere valutate, migliorate e apprese (vedi paragrafo 1.2). Oltre all'aspetto formativo, la vera innovazione di ENACT è rappresentata dalla possibilità di valutare in maniera diretta le competenze oggetto di studio attraverso opportune metodologie quantitative, aspetto poco o per niente esplorato in questo campo.

ENACT - Enhancing Negotiation skills through on-line Assessment of Competencies and interactive mobile Training, è un progetto finanziato dalla Comunità Europea nell'ambito del Programma Lifelong learning, che ha coinvolto nel biennio 2014-2015 sei partner provenienti da 4 paesi europei. Il progetto è stato coordinato dalla Plymouth University (UK) in collaborazione con Aidvanced (Italia), Università di Napoli "Federico II" (Italia), Fondazione Mondo Digitale (Italia), Ministero turco per la gioventù e lo sport (Turchia) e Fundetec (Spagna). Nel corso del progetto, il consorzio ha lavorato allo sviluppo dello strumento, sia sul piano tecnologico che contenutistico, e ha effettuato una serie di trial sperimentali per verificarne la validità scientifica. Ciascun partner ha contribuito alla realizzazione e alla validazione dello strumento in momenti diversi e in base alle proprie competenze specifiche. Il NAC ha fornito contributi relativi all'analisi dei bisogni degli utenti, alla creazione dei testi implementati e alla realizzazione dei trial sperimentali.

3.2.1 Obiettivi del progetto

L'obiettivo principale di ENACT è la realizzazione di un metodo di insegnamento e di valutazione, innovativo, flessibile, e gratuito, per migliorare la negoziazione e

capacità di comunicazione in contesti educativi formali e informali. A tal fine sono stati progettati e realizzati due strumenti specifici:

- 1) **ENACT Training:** metodologia di formazione innovativa e diversificata, in forma di "gioco" on-line, disponibile sia in versione guidata che di auto-apprendimento.
- 2) **ENACT Assessment:** metodologia di valutazione integrata in un serious games (SGS), in grado di fornire un'analisi attendibile delle capacità di gestione dei conflitti e sugli stili comunicativi dell'utente.

La Figura 5 mostra le due versioni del gioco disponibili online all'indirizzo www.enactskills.eu.



Figura 5: ENACT game

La presente ricerca si focalizza sulla seconda applicazione, ENACT come strumento psicodiagnostico. Come già detto in precedenza, l'assessment in ENACT è stata realizzata implementando il modello dei 5 fattori di gestione del conflitto di Rahim. Di fatto questa implementazione si configura come una traslazione del test ROCI II, in una serie di simulazione di conflitti virtuali tra un utente e il sistema, in un ambiente 3D, al fine di valutare le competenze specifiche del soggetto. L'obiettivo è verificare la validità e l'attendibilità di ENACT mediante uno studio di comparazione psicometrica tra il nuovo strumento e il test di Rahim.

3.2.2 ENACT Assessment: lo strumento

ENACT Assessment (da ora solo ENACT) è un serious game basato su una piattaforma web, sviluppata utilizzando il framework Unity (<https://unity3d.com/>), realizzato in **due versioni**, una per ogni genere e in **4 lingue**: inglese, italiano, spagnolo e turco. **Il gioco è organizzato in scenari “di conflitto” indipendenti, in ciascuno dei quali l'utente ha il compito di impersonificare un diverso personaggio e negoziare con diversi agenti artificiali (bot) in situazioni di vita realistiche e quotidiane.** In ogni scenario la negoziazione è resa attraverso una serie di interazioni uomo/bot, in ciascuna delle quali il soggetto ha la possibilità di scegliere una frase in un gruppo di 5, per giungere ad una risoluzione del conflitto. Ciascuna frase è indicativa di uno degli stili identificati da Rahim. Gli scenari sono bilanciati rispetto al genere del bot e l'ordine delle frasi è casuale. Il software consente di raccogliere sia le risposte che i tempi di interscambio dell'utente. Alla fine della sessione, viene fornita una scheda di valutazione del comportamento adottato in tutti gli scenari e una suggerimento per migliorare la propria performance. L'intera sessione di gioco è visibile nell'allegato 2.

ENACT è uno strumento pensato per raggiungere tutti coloro, individui e organizzazioni, che hanno un accesso limitato alla formazione e alla valutazione in questo settore a causa dei costi spesso inaccessibili e/o per vincoli organizzativi. Risponde molto bene a questa esigenza anche perché è un sistema **asincrono** che non richiede cioè la partecipazione simultanea di più gli utenti, permettendo allo stesso tempo di esercitarsi in un ambiente “sicuro” e altamente motivante.

L'avvio del gioco prevede la registrazione dell'utente, al fine di personalizzare il gioco in base a parametri individuali quali il genere e la nazionalità. In ogni sessione di gioco, per ogni scenario, due schermate introducono la situazione conflittuale, i personaggi coinvolti, il ruolo assegnato all'utente e il suo obiettivo finale. Vengono fornite inoltre anche informazioni sul tempo a disposizione e sull'importanza data da entrambi i protagonisti al raggiungimento dello scopo.

Come mostra la Figura 6, l'interfaccia è stata progettata per essere il più intuitiva possibile, richiedendo all'utente di effettuare solo azioni semplici.



Figura 6: Istruzioni di gioco in uno scenario ENACT

All'interno di ogni scenario, entrambi i personaggi sono rappresentati da un avatar, di cui uno è controllato dal computer ed è posizionato nella parte centrale dello schermo, mentre il secondo è controllato dall'utente ed è posizionato in una finestra più piccola nell'angolo in alto a sinistra (Figura 7). Entrambi gli avatar sono in grado di trasmettere emozioni e azioni attraverso animazioni facciali e posturali. Come mostra la figura, le frasi a sinistra rappresentano le scelte del giocatore in risposta alle azioni del bot. Spostando il mouse sulle singole frasi, l'avatar del giocatore si muove, assumendo espressioni ed effettuando gesti correlati al significato della frase. L'utente effettua la propria scelta solo cliccando sulla frase che ritiene più adatta alla situazione, ripetendo l'azione per 4 interazione in base alla risposta del bot.



Figura 7: Esempio di interazione in uno scenario di ENACT

L'ultima schermata di ciascuno scenario rappresenta una domanda finale, che può essere considerato un item di controllo (Figura 8), con la quale si richiede al soggetto di scegliere un'azione tra le cinque possibili, ognuna delle quali è costruita in base ad uno specifico stile del modello di Rahim. Questa variabile ha implicazioni importanti nell'analisi dei dati, come dimostrato nelle sezioni successive (paragrafo 3.2.6.1).

Alla fine dello scenario, l'utente non riceve alcun feedback, ma alla fine della sessione di valutazione viene fornita una scheda di debriefing, con un'analisi degli stili di negoziazione utilizzati nei 5 scenari (sulla base del modello di Rahim) e un suggerimento per migliorare la propria prestazione (Figura 9). L'obiettivo è quello di restituire una valutazione paragonabile a quella ottenuta con il test ROCI II, finalizzata all'elaborazione automatica di una strategia di formazione personalizzata e centrata sullo studente, da utilizzare in ENACT Training.



Figura 8: Schermata finale di una sessione di assessment



Figura 9: Scheda di valutazione finale di ENACT Assessment

3.2.3 Le caratteristiche principali di ENACT

Come abbiamo detto, ENACT implementa il modello teorico di Rahim sui 5 fattori di gestione del conflitto. Tale implementazione è stato frutto di un lungo lavoro finalizzato all'identificazione di una serie di elementi che di fatto l'hanno reso possibile. Innanzitutto, il gioco utilizza **avatar animati settati su una serie di espressioni di base**, sia di tipo verbale, come il tono vocale e la struttura della frase, che non verbale, come le espressioni del viso, il contatto visivo, la postura del corpo e i gesti (vedi Figura 10).



Figura 10: Esempi di espressioni e gesti degli avatar di ENACT

Queste espressioni sono rese dalla combinazione di **7 indicatori specifici dei diversi stili di comunicazione**, rilevatori di altrettanti tratti comportamentali osservabili e misurabili:

1) Struttura della frase: resa sempre in prima persona con una chiara indicazione delle intenzioni del soggetto, operazionalizzata nelle frasi come ad esempio "io preferirei", " Io non sono sicuro...", ecc.

2) Fatti-opinioni: la distinzione chiara nelle frasi dei personaggi tra opinioni e fatti oggettivi, ad esempio, "Penso che", "È tutta colpa tua", "Potrei sbagliarmi";

3) Richiesta: vale a dire, l'inserimento di elementi situazionali, di tipo inclusivo e aperto al discussione, esemplificati da frasi quali "Come ti senti a riguardo?", o "Non mi importa quello che pensi, basta farlo!", che invece ha carattere opposto;

4) Critica: misura di un atteggiamento critico costruttivo contro uno manipolativo o auto-distruttivo, come ad esempio "Potresti trovarlo più utile se" contro "Se fossi in te farei" oppure "Io sono completamente inutile nel fare;"

5) Tono vocale: rappresentato graficamente dalla forma della vignetta

6) Espressioni facciali: rappresentati dalle animazioni 3D degli avatar, presenti anche nei periodi di inattività in fase di attesa.

7) Gesti corporei: rappresentati dalle animazioni 3D che si attivano mentre gli avatar pronunciano le diverse frasi.

L'efficacia e la univocità di tali variabili e le loro implementazioni grafiche sono state valutate nei test di usabilità, realizzati all'inizio del progetto, analizzando i dati raccolti con questionari post-test ai diversi gruppi di utenti (per una spiegazione più dettagliata vedere paragrafo 3.2.5). Lo sviluppo di avatar animati ha permesso **l'implementazione del modello di Rahim**, sia attraverso la caratterizzazione delle frasi utilizzate negli scambi, sia dando ai bot una propria "personalità", che emerge durante le interazioni in base alla situazione di conflitto proposta. Ogni bot ha un comportamento statico, non adattivo che riflette uno specifico stile di negoziazione. Le azioni sono calcolate in base al modello psicologico incorporato e le caratteristiche specifiche associate ad esso all'interno dello scenario. Questo significa che un bot dominante mostrerà comportamenti prevalentemente aggressivi e

autorevoli. Al contrario, un bot servizievole mostrerà un atteggiamento passivo e generalmente sottomesso. Inoltre, per controllare l'effetto di possibili influenze culturali, il sistema prevede avatar di entrambi i generi, con un'età compresa tra i 20-40 anni, appartenenti a diversi gruppi etnici: europei, asiatici, africani e arabi. L'attribuzione del **genere del personaggio interpretato** è un altro elemento peculiare di ENACT. Il sistema è infatti settato per attribuire all'utente sempre personaggi del suo stesso sesso (informazione questa ottenuta in fase di registrazione). È evidente che questo aspetto ha forti ricadute sui testi della conversazione che devono prevedere entrambe le tipologie di scambio (maschio-femmina; femmina-maschio). Ma il vero aspetto innovativo del gioco è **l'assessment**, che implementa la misurazione delle capacità di negoziare con metodologie innovative e con un approccio psicometrico rigoroso. Il sistema raccoglie i dati sulle scelte degli utenti in relazione al comportamento del bot e le scelte effettuate in ciascuna interazione al fine di creare un modello del giocatore. Il profilo è dato dalla somma dei punteggi ottenuti dai soggetti sui 5 fattori di Rahim, rappresentati negli scenari dalle frasi in ciascuna interazione, che permette poi di definire il profilo dell'utente a seconda di come esso si posiziona sullo schema del modello di Rahim, vedi Figura 4. Questo è reso possibile dall'implementazione di un **sistema di tutoring** interno al software che alla fine della sessione di valutazione crea automaticamente un feedback per l'utente sui punti di forza e le aree di miglioramento.

3.2.4 La validazione

L'obiettivo finale di ENACT è quello di restituire ai partecipanti una valutazione degli stili di gestione dei conflitti paragonabile a quelle fornite da test carta e matita, ROCI II. A tal fine, nel corso del progetto, lo strumento è stato sottoposto ad un processo di valutazione, articolato in più fasi (vedi paragrafo 3.2.5) volto a finalizzare le diverse componenti del nuovo tool. Per la validazione psicometrica della prima versione del software è stata utilizzata una batteria di 5 test standardizzati e normati, per i tre diversi paesi, Italia, Turchia, Spagna, composta da:

1) **ROCI II - Modulo C**, test di **28 item** misurati su scala likert, specifico per i conflitti con i pari¹², descritto nel capitolo precedente.

2) **BFA** (Big Five Adjectives): test breve di personalità, basato sul **BIG FIVE**, uno dei più noti test sulla personalità, che individua 5 dimensioni di base o tratti della personalità (modello a cinque fattori (FFM) di Barbaranelli e Caprara, 2000; 2002), quali: apertura mentale, coscienziosità, estroversione, gradevolezza e nevroticismo. Il test presenta **20 aggettivi**, che descrivono i cinque tratti della personalità e i soggetti devono scegliere il loro grado di accordo con ogni aggettivo.

3 e 4) **Negotiation and Assertive efficacy**: due misure specifiche del concetto più ampio, sviluppato da Bandura (1994) di **autoefficacia percepita**¹³. Nello specifico, la **Negotiation efficacy** può essere definita come la **fiducia di ciascuno nella proprie capacità di applicare strategie di negoziazione** adatte a superare una situazione difficile; mentre **Assertive efficacy** si riferisce la **fiducia che ciascuno ha in se stesso** in una data situazione e sono misurati rispettivamente con una **scala a 8 e 6 item**.

5) **Coping**: strettamente legato allo stress, il coping è **definito come un processo dinamico tra una persona e l'ambiente circostante** (Teoria transazionale di stress e di coping di Lazarus e Folkman, 1984). Folkman e colleghi (1986) hanno trovato **8 diverse strategie di coping** che possono essere utilizzate per reagire a situazioni stressanti. Le Strategie di coping sono state misurate con una versione abbreviata del test «**Ways of Coping**», originariamente sviluppato da Folkman e Lazarus (1980; 1988), con otto **8 sub scale** che misurano le **8 strategie** individuate dagli autori.

La scelta di ognuno di questi test è finalizzata alla verifica della validità concorrente e discriminante di ENACT, rispetto a diversi fattori. Con il Big Five, ad esempio, ci si propone di verificare se un modello di personalità ha un impatto diretto sullo stile di gestione utilizzato, le scale Negotiation and Assertive Efficacy mirano a valutare

¹², i moduli A e B sono orientati a misurare stili di gestione di conflitti rispettivamente con i superiori e subordinati.

¹³ Per Bandura, un personale senso di efficacia rappresenta la base dell'agire umano (Bandura, Pastorelli, Barbaranelli, e Caprara, 1999). Inoltre, molte ricerche dimostrano il ruolo importante di autoefficacia nel regolare funzionamento umano e il benessere per mezzo di processi cognitivi, emotivi, affettivi e selettivi (Bandura, 1997).

la percezione personale sulla propria capacità di affrontare e gestire situazioni difficili, come ad esempio situazioni di conflitto, mentre la scala di Coping valuta le modalità dell'individuo di reagire a situazioni stressanti, come appunto può essere un conflitto.

3.2.5 L'Implementazione

Dopo una prima fase di progettazione e sviluppo, ENACT è stato sottoposto a tre cicli di valutazione e revisione, finalizzati alla realizzazione dello strumento così come è stato descritto nel paragrafo 3.2.2. Ogni ciclo era mirato allo studio e al miglioramento di uno o più elementi del sistema.

1) Test di usabilità.

Al fine di valutarne l'usabilità, la versione beta del serious game ha subito due iterazioni complete di sviluppo e, dopo ogni iterazione, il software è stato testato per ottenere informazioni su elementi incisivi per l'efficacia e l'accettabilità della piattaforma. Il 1° test è stato effettuato presso lo Showcase SciTech 2014, alla Plymouth University, nel settembre 2014. La prima demo del ENACT (1 solo scenario) è stato presentato ad un gruppo di 79 soggetti di età compresa tra gli 11 e i 60 anni, (età media 20,6, 41 maschi e 38 femmine). I soggetti hanno risposto ad un questionario di 7 domande inerenti al contenuto del gioco e all'interfaccia, misurati su scala Likert a 5 punti, e un ottavo item, dicotomico, sulla volontà di rigiocare. I dati hanno mostrato un feedback generale positivo (media 3.5). Gli aspetti più criticati riguardavano il realismo delle conversazioni e la grafica del gioco. Tuttavia, il 95,3% dei soggetti ha mostrato un interesse a ripetere l'esperienza. A seguito di questa prima valutazione, il software è stato modificato nei contenuti, in alcuni aspetti della grafica ed è stato ampliato. La seconda versione è stata sottoposta a valutazione durante la British Science Week ospitata dall'Università di Plymouth nel marzo 2015. In questo occasione, la versione migliorata del software si componeva di 4 scenari differenti e una valutazione della performance. Il campione era composto da 39 persone, di età compresa tra 11- 28 anni, di cui 18 maschi e 21 femmine. Il questionario era composto da 13 domande (5 in più rispetto al precedente), di cui 12 su scala Likert a 5 punti, e una dicotomica sull'intenzione a rigiocare. L'analisi dei dati ha mostrato un atteggiamento molto positivo nei confronti del gioco,

evidenziando un leggero miglioramento rispetto al test precedente (nessun punteggio al di sotto di 3.59). Anche in questo caso, quasi il 95% dei soggetti ha sostenuto l'intenzione a riprovare il test. Dopo queste due fasi di valutazione si è proceduto alla realizzazione della **1° versione di ENACT ASSESSMENT, utilizzato nel test pilota.**

2) Test pilota

Il test pilota è stato condotto utilizzando la prima versione del software con l'obiettivo di tarare la metodologia e il sistema di valutazione e migliorare l'usabilità dello strumento. A tal fine, ENACT, insieme alla batteria di 5 test sopra descritti (3.2.4), è stato somministrato ad un campione di 146 soggetti provenienti da tre diversi paesi. In particolare, sono stati coinvolti 96 maschi e 50 femmine, di età compresa tra 14 e 63 anni (media 27.34 e deviazione standard 8.93) e 5 gruppi target: studenti delle scuole superiori (Mondo Digitale, Italia), ricercatori in formazione (Università di Napoli Federico II, Italia), imprenditori delle PMI (Fundetec, Spagna), studenti universitari (Università di Plymouth, Regno Unito) e atleti (Ministero della Gioventù e dello Sport, Turchia). I testi del gioco sono stati tradotti dall'Inglese all'italiano, spagnolo e turco utilizzando il protocollo di back traslation. I dati sono stati raccolti elettronicamente attraverso un server proprietario, per ENACT, e il software Modulo di Google per la batteria di test di controllo.

È importante sottolineare che in questa prima versione, è stato deciso di escludere lo stile Compromising dagli stili di negoziazione degli avatar. Questa scelta, di stampo puramente metodologico, muove da un'analisi dello stesso modello di Rahim, che pone questo stile al centro degli altri 4 (vedi Figura 4) non trovando pertanto una pura identificazione fattoriale.

I risultati dello studio pilota sono stati incoraggianti. Per l'attendibilità si è analizzato l'Alpha di Cronbach delle diverse scale, con un buon risultato sulle dimensioni Dominating e Integrating, entrambe maggiore di 0.8, Obliging supera di poco 0.6 mentre Avoiding appare poco attendibile. Un'analisi più approfondita dell'attendibilità di ciascuno stile nei singoli scenari, mostra nelle ultime due scale, problemi di coerenza tra gli item (scambi comunicativi). Anche gli studi sulla validità hanno mostrato un risultato discreto. La validità discriminante è stata dimostrata dall'assenza di correlazioni tra tratti di personalità ed ENACT, come ci si

aspettava, dato che ENACT non è pensato per misurare questi tratti. Mentre la validità concorrente è dimostrata, parzialmente, da correlazioni statisticamente significative per le scale Dominating e Integrating di ENACT con quelle del ROCI, così come da una serie di buone correlazione tra ENACT e le strategie di Coping. In particolare, si riscontra una significativa correlazione positiva tra la scala confrontive (relativa all'ostilità e all'assunzione di rischi), e lo stile Dominating. È interessante notare che, utilizzare uno stile dominante correla positivamente con escape-avoiding coping strategy, suggerendo che mediare con un avversario può essere un'attività stressante. Inoltre, le significative correlazioni positive tra le strategie di Coping: ricerca di supporti sociali e accettazione delle responsabilità, e lo stile Integrating mostra un altro punto a favore della validità convergente: ossia accettare la responsabilità e la ricerca di sostegno sociale svolgono un ruolo importante per l'integrazione degli interessi delle parti in conflitto. Infine, non risultano statisticamente significative invece i) l'ordine di somministrazione dei test, gestito in maniera random; ii) il genere del Bot e iii) il posizionamento delle frasi.

I risultati ottenuti con lo studio pilota sono stati interessanti e sicuramente incoraggianti, ma hanno rappresentato solo il primo passo nel processo di validazione di ENACT come strumento di misura degli stili di negoziazione. L'attendibilità e la validità dello strumento sono parziali, dimostrando problematiche con almeno due scale di misura (Obliging and Avoiding), senza contare che questa prima versione mancava di fatto della scala Compromising. In generale, le principali problematiche sembrano essere legate agli scenari utilizzati e alla poca corrispondenza con gli item originali.

3) Trial Sperimentale

In seguito ai risultati ottenuti con il test pilota, gli sviluppatori hanno deciso di apportare una serie di cambiamenti al software con l'obiettivo di rendere lo strumento più valido e affidabile. In particolare: 1) sono stati arricchiti le descrizioni dei diversi scenari in modo da consentire agli utenti di comprendere meglio l'importanza del conflitto; e 2) è stato aggiunto lo stile Compromising, non considerato nella prima versione, al fine di mantenere una maggiore coerenza con lo strumento originale; 3) sono stati apportati importanti miglioramenti al layout e alla

grafica. La versione finale, così come descritta nel paragrafo 3.2.3, è stata utilizzata in un trial sperimentale, che ha coinvolto 130 soggetti, 104 maschi e 26 femmine, di età compresa tra i 13 e i 60 anni (media 21.44 e deviazione standard 93.72), e divisi in 5 gruppi target, esattamente come nel test pilota: studenti delle scuole superiori (Italia), i ricercatori in formazione (Italia), imprenditori delle PMI (Spagna), studenti universitari (UK) e atleti (Turchia). Anche in questo caso i nuovi testi sono stati tradotti dall'inglese all'italiano, spagnolo e turco utilizzando il protocollo di traduzione back traslation. Dati i risultati ottenuti nel test pilota, in questo ciclo non è stata somministrata ai partecipanti l'intera batteria di controllo, ma solo il ROCI II. Anche in questo caso i dati sono stati raccolti automaticamente con le stesse modalità del test pilota. Per quanto riguarda i **risultati**, sono state svolte analisi di attendibilità e validità sul campione di dati utilizzando l'alfa di Cronbach e la correlazione tra le diverse scale di ENACT con quelle di ROCI II. I primi risultati possono essere interpretati positivamente. Infatti, anche se l'affidabilità in termini di Alfa Cronbach sono inferiori al quelli della fase precedente, si osserva una significativa correlazione tra lo stile Obliging misurato da Enact e quella di Rahim (non presente nella versione precedente). La nuova versione ha portato da due a tre gli stili misurati con affidabilità e validità. È interessante notare che le correlazioni calcolate in ciascuno scenario, utilizzando l'ultima risposta o item di controllo, hanno mostrato risultati simili. In particolare, gli stili Integrating, Dominating e Obliging correlano significativamente con le loro controparti del ROCI (0,58, 0,47 e 0,42, rispettivamente). I problemi relativi agli stili Avoiding and Compromising possono essere legati al fatto che essi presentano una correlazione evidente con altri due stili: Obliging and Integrating. Infine, le correlazioni sono statisticamente significative ma non molto elevate. Questo in parte si spiega con la grande differenza tra i due strumenti, sia in termini di ricchezza dei dati che di modalità di somministrazione.

3.2.6 ENACT – ROCI II: Discussione Critica Dei Risultati

A conclusione del progetto ENACT, i risultati possono essere interpretati in maniera positiva. Sebbene le analisi mostrano infatti alcune criticità, soprattutto su alcune delle scale investigate, è pur vero che ENACT è il primo tentativo di realizzare un test psicologico con la tecnica del serious game. Come si evince dall'analisi della letteratura, creare e validare un test psicologico è già di per se un'attività molto impegnativa che richiede un gran numero di partecipanti e un adeguato numero di sperimentatori, con una buona dose di pazienza. Nel caso di ENACT, a questi elementi vanno aggiunti un set di criticità legate allo sviluppo di una piattaforma software e tutti i problemi tecnici che questo comporta: progettazione 3D, raccolta dei dati, usabilità e design, solo per citarne alcuni. Nonostante le difficoltà, si ritiene che lo sviluppo di test sotto forma di giochi interattivi rappresenti il futuro della valutazione psicologica. I serious games hanno un potenziale non ancora del tutto evidente in questo settore. Secondo Nielsen, gli americani a partire dai 13 anni hanno trascorso nel 2013, n. 6,3 ore a settimana a giocare con i videogame, su qualsiasi piattaforma (5,6 nel 2012 e 5,1 nel 2011). Una quantità enorme di tempo che può essere parzialmente sfruttato per la somministrazione di test psicologici, in grado di fornire agli individui un feedback importante su se stessi, e forme innovative di formazione.

Quello che emerge, da questi primi risultati, riguarda soprattutto le criticità legate alle differenze tra i due strumenti e le difficoltà di confrontare i dati raccolti.

Come già abbiamo detto, il ROCI è un test carta e matita, composto da 28 domande misurate su scala Likert, ognuna delle quali chiede di richiamare alla memoria situazioni di conflitti vissute dal soggetto negli ultimi mesi, con i propri pari, dando una valutazione che va dal "Per niente d'accordo" al "Completamente D'accordo". Il test si compone di 5 gruppi di item, ognuno predisposto a misurare uno specifico stile di risoluzione del conflitto. L'analisi fattoriale di Rahim mostra come gli item si raggruppano in maniera convincente e salda sui 5 fattori da lui identificati. I dati risultano quindi quantitativi, indipendenti e chiari.

ENACT è un software di gioco, basato sulla tecnica del role play, composto da 5 scenari che propongono 5 diverse situazioni di conflitto relativi ad argomenti di vita

quotidiana. In ogni scenario, l'utente interagisce con un bot, caratterizzato da un proprio stile di negoziazione, e ha a disposizione 5 scambi comunicativi per raggiungere la soluzione più adatta al problema. Ad ogni scelta dell'utente, il bot utilizza una risposta tra le 5 possibili, che a sua volta determina l'estrazione, da parte del sistema, di un pattern di item da proporre all'utente nello scambio successivo, e così via. Questo schema, così complesso, si configura come un albero, dove ogni scambio rappresenta una diversa ramificazione, e dove quindi ogni utente alla fine sarà caratterizzato da un proprio profilo specifico. A questo si aggiungono alcune variabili importanti, che il test proposto da Rahim, data la sua natura, non può tenere conto, quali il tempo utile per la risoluzione del conflitto e l'importanza per il soggetto di giungere ad una soddisfazione piena, l'ottenimento parziale o completo del proprio obiettivo, sebbene l'autore li ritiene elementi importanti nel suo modello.

Quindi, sebbene partano dagli stessi assunti teorici, ENACT infatti nasce con l'idea di rappresentare una forma parallela del ROCI, i due strumenti sono caratterizzati da profonde differenze, che influenzano di fatto l'analisi dei dati e di conseguenza i risultati.

3.2.6.1 Le differenze e la loro influenza nell'analisi dei dati

La prima differenza significativa, che si riflette sull'analisi dei dati, è che ENACT **non richiama eventi di vita vissuta**. In quanto serious game basato sul role-play, simula situazioni di vita e spinge l'utente a immedesimarsi in uno dei protagonisti, indipendentemente se la situazione è stata vissuta oppure no. Questo implica che il soggetto potrebbe anche non aver mai vissuto esperienze simili e non potendo far ricorso alla sua esperienza, risponde al test in base al suo immaginario, che non è stato mai testato dall'utente precedentemente, oppure fa ricorso a situazioni vicine ma non necessariamente uguali, dove anche un solo cambiamento (intesa come variabile) può determinare grandi differenze (ad es. fa ricorso al ricordo di un conflitto vissuto con un superiore, invece che con un pari).

La seconda differenza è data dal diverso **stile comunicativo**. Essendo un test carta e matita, il ROCI utilizza uno stilo comunicativo chiaro e inequivocabile. Il linguaggio utilizzato permette di differenziare bene tra le diverse modalità di risoluzione del conflitto, attraverso l'uso di parole chiave per definire esattamente il tipo di ricordo

da esplicitare. A titolo esemplificativo si riportano alcuni esempi di item del ROCI II (per il test completo vedi l'allegato 1):

- Ñ (1 INT) Cerco di indagare su un problema con i miei colleghi per trovare una soluzione accettabile per tutti.
- Ñ (15 COM) Negoziare con i miei coetanei in modo che un si possa trovare un compromesso
- Ñ (3 AVO). Cerco di evitare di essere "messo in difficoltà" e tendo a tenere il mio conflitto per me.
- Ñ (11 OBL). Generalmente cerco di soddisfare le esigenze dei miei coetanei.
- Ñ (9 DOM). Uso la mia autorità per prendere decisioni a mio favore.

ENACT basandosi su uno scambio comunicativo di tipo colloquiale, non può utilizzare parole “chiave” come “soluzione integrata” o “compromesso”, ma deve elicitarle tali concetti in maniera implicita, e questo porta alle difficoltà di differenziare le 5 opzioni di risposta, soprattutto per gli stili molto simili, come Integrating e Compromising.

La terza differenza è data dalla **scala di misura**. Il Roci utilizza una scala Likert a 5 passi, che nella teoria classica dei test è assimilata ad una scala di misura quantitativa. Con ENACT si ottengono dati di tipo qualitativo, poiché per ogni scenario proposto si raccolgono delle frequenze, ossia il numero di scelte fatte dal soggetto per ciascuno stile. Nelle analisi di fine progetto questa difficoltà è stata superata trattando i dati di ENACT come misure quantitative, in base ad una serie di considerazioni relative alla costruzione degli scenari. Più nello specifico, si è ritenuto che ogni risposta, delle 5 possibili, rappresenti un punteggio sulla scala di quello specifico stile, e quindi la somma delle risposte dello stesso stile per ciascuno scenario può essere considerato come una misura quantitativa. Più specificatamente, le considerazioni alla base di questo metodo sono che: i) il ROCI non ha scenari, ma tiene conto solo degli stili, per questo ii) il singolo scenario può essere considerato un item “a più risposte” dove iii) più volte uno stile viene scelto dal soggetto in uno specifico scenario, tanto più è alto il punteggio ottenuto dal soggetto per quello stile su una scala di misura simile a quella Likert utilizzata per il ROCI II.

La conseguenza diretta di questa metodologia di analisi è che non sia ha certezza sui risultati ottenuti e che, quantomeno, queste considerazioni andrebbero verificate con un metodo statisticamente convincente.

Un'altra differenza significativa è rappresentato dall'**Item n.5** di ciascuno scenario di ENACT. Esso, infatti, a differenza dei precedenti, si configura come una sintesi del comportamento, una misura di controllo della coerenza dello scenario. Quindi, sebbene consecutivo e logicamente connesso ai precedenti, ha una natura completamente diversa, potremmo definirlo una sorta di **item di controllo**. Per meglio comprendere la differenza tra la risposta 5 e le precedenti, la tabella che segue mostra un esempio tratto dallo scenario "Programmi TV".

	Player	Bot (stile Dominating)
Introduzione		Che sfortuna! Stasera trasmetteranno sia Masterchef che Report!
	Quale programma guardiamo?	Semplice, si guarda Masterchef senza discutere!
Interazione 1	Che ne dici invece di trovare una soluzione che faccia piacere ad entrambi? (Integrating)	
		E come? Non esiste! Non ci provare.
Interazione 2	Ma come mai hanno messo i due programmi alla stessa ora? (Avoiding)	
		Stasera sono stressata e voglio vedere Masterchef per rilassarmi. Report lo replicano domani!
Interazione 3	Se per te Masterchef è così importante io potrei vedere la replica di Report...(Obliging)	
		Hai indovinato! Stasera guardo io Masterchef!
Interazione 4	Non posso vederti così stressata, stasera Masterchef! Tu sei la mia priorità. (Obliging)	
		Meglio se cambi stanza, così forse riesco a rilassarmi un po'.
ITEM 5 In sintesi.. In base a quanto accaduto finora, cosa pensi di fare ora?"	Proverò nuovamente a proporre a [...] di vedere le parti più interessanti di entrambi i programmi. (Integrating)	
	Rimarrò saldo sulla mia posizione: voglio vedere Report. (Dominating)	
	Come d'accordo con [...], vedremo Masterchef. (Obliging)	
	Proverò a rimandare la decisione a dopo che avrò finito di sistemare i documenti.	

	(Avoiding)	
	Proverò nuovamente a proporre a [...] di vedere entrambi i programmi, ciascuno per metà del tempo.	

Tabella 6: Differenza degli item in un scenario ENACT

È facile notare che il soggetto, di fronte ad un bot molto determinato (settato sullo stile Dominating), dapprima ha cercato una soluzione integrata, dopo ha cercato di evitare la questione e alla fine si è arreso lasciando la decisione all'interlocutore (Obliging) per ben due volte. Questa stessa decisione è stata poi confermata anche dalla risposta alla domanda finale. È facile notare in questo esempio la differenza tra i due tipi di item proposti. Per questa differenza, la risposta numero 5 non è stata considerata nell'analisi sulla dipendenza descritte nel paragrafo 3.3.3, ma è stato necessario analizzarla singolarmente. È importante sottolineare che l'inserimento di questo tipo di item è stato dettato inizialmente solo dalla necessità di creare una conclusione alla sessione di gioco, e solo in un secondo momento, e cioè nella fase di analisi esplorativa, descritta nei paragrafi successivi, ci si è resi conto delle differenze interne e del loro impatto sulle analisi. In conclusione, quello che emerge, è che non solo l'item n. 5 è differente dai primi 4, ma che non esiste un item simile nel test di ROCI, rendendo i due test, da questo punto di vista, non confrontabili direttamente.

Infine, la quinta differenza è l'**indipendenza dei dati**. Nel Roci, infatti, ogni item è indipendente dagli altri. L'analisi fattoriale di Rahim mostra come ogni item misura e spiega una parte specifica dei 5 fattori estratti, uno per ogni stili di risoluzione del conflitto. Questo implica che la risposta ad un item non ha alcuna influenza sulle risposte agli altri, anche se appartenenti alla stessa scala. Questo non è vero invece in ENACT, dove ogni singola risposta influenza il comportamento del BOT, e quindi la sua risposta a quel determinato scambio, che a sua volta determina la scelta del pattern di item proposto all'utente nello scambio successivo, e così via, creando di fatto un rapporto di dipendenza, quantomeno teorico, tra gli item dello stesso scenario.

3.2.6.2 Le questioni aperte

Ciascuna di queste differenze sollevano una o più domande sulla bontà dei dati e di conseguenza sui risultati ottenuti. In particolare:

- Il tipo di dato ottenuto con ENACT ha lo stesso valore e significato di quello ottenuto con il ROCI? Se sì, i dati sono quindi confrontabili in maniera diretta? In caso contrario, come bisogna analizzare i dati di ENACT per renderli confrontabili con quelli di Rahim?
- Le tecniche classiche della psicometria, utilizzate per Rahim, riescono a cogliere e misurare quello che si intende misurare con ENACT?
- E più in generale, possono queste tecniche “tradizionali” essere applicate ad uno strumento così peculiare e innovativo come ENACT, senza una perdita o una distorsione delle informazioni?
- È possibile pensare una tecnica o un procedimento che rende questi due tipi di strumenti confrontabili in maniera valida e attendibile?

Per cercare di dare risposta a queste domande, questo studio propone **un’analisi esplorativa** dei dati raccolti nel progetto, ossia *un’analisi finalizzata alla ricerca di una serie informazioni, precedentemente sconosciute e potenzialmente utili, all’interno di grandi basi di dati*. L’obiettivo finale è l’individuazione di una struttura, un modello, o, in generale una rappresentazione sintetica dei dati, in grado da una parte di comparare i due strumenti e dall’altra di definire le principali criticità offrendo indicazioni utili affinché queste siano superate negli studi successivi.

Con questa analisi, in sostanza, si ricerca un protocollo di misura capace di chiarire la natura dei dati oggetto di studio e dall’altra di identificare una tecnica “tradizionale”, in grado di comparare i due test. Si cerca dunque un protocollo di analisi che tenga conto di questi elementi distintivi e permetta di ottenere dati confrontabili validi. Nel capitolo che segue si descrive la natura e la metodologia utilizzata per effettuare questo confronto.

3.3 Analisi esplorativa sui risultati del progetto

Nell'ambito del progetto, i dati ottenuti dalle sperimentazioni sono stati analizzati con le tecniche classiche dell'analisi dei dati normalmente utilizzate per i test "carta e matita", volte a misurare l'attendibilità e la validità dello strumento in relazione al test di origine, il ROCI II. Queste analisi se da una parte hanno evidenziato degli aspetti positivi del serious games, dall'altra hanno lasciato una serie di quesiti aperti, legati soprattutto alla possibilità di confrontare i due strumenti.

La grande differenza dei dati prodotti dai due test ci ha indotto ad investigare sulla natura del test realizzato e verificare le possibilità di confronto con il ROCI. Per questo motivo, l'ultima fase del presente progetto di ricerca si è incentrata su un'Analisi dei Dati Esplorativa (ADE), ossia un'analisi statistica volto a riassumere e identificare le caratteristiche principali di una base di dati molto grande, spesso utilizzando metodi grafici. In altre parole, con l'ADE, si effettua una esplorazione dei dati priva di ogni modellizzazione o ipotesi inferenziale, per scoprire quello che i dati possono dire sulla loro natura, al fine di, eventualmente, formulare nuove ipotesi che potrebbero portare ad nuova organizzazione e sperimentazione. Si sono utilizzate pertanto una serie di tecniche di analisi, descritte nel dettaglio nel paragrafo che segue, con l'obiettivo principale di individuare attraverso l'analisi dei dati, un metodo statistico valido che ci permetta di confrontare i due strumenti, metodo che possa essere usato come riferimento in futuro in questo tipo di comparazione e che soprattutto permetta di individuare gli elementi positivi e le criticità di cui tener conto nei successivi studi.

Partendo dalle questioni aperte descritte nel paragrafo precedente, i dati raccolti nello studio pilota sono stati pertanto trattati e analizzati per rispondere ai seguenti quesiti:

- *I punteggi ottenuti dai soggetti nei singoli scenari di ENACT possono essere considerati misure indipendenti oppure data la struttura del test, i punteggi devono tener conto della consequenzialità delle risposte e quindi devono essere considerate dipendenti tra loro?*
- *I 4 item di ciascuno scenario sono coerenti con l'item di controllo?*

- *Gli scenari sono caratterizzati da uno stile di risposta predominante o sono stati costruiti senza influenze stilistiche?*
- *Gli stili misurati dai due strumenti e più in generali i due strumenti stessi hanno una struttura latente simile? Misurano gli stessi concetti? Come possono essere confrontati?*

Lo studio è stato articolato in quattro fasi principali, ognuna caratterizzata dall'utilizzo di specifiche tecniche statistiche e ognuna volta a rispondere ad una di queste domande. L'intera analisi si apre con una fase di preparazione dei dati, fase necessaria data la complessità e la ricchezza di informazione presente nel database di raccolta dati proprio del software ENACT.

3.3.1 Metodologia d'analisi

L'Analisi esplorativa dei dati, detta appunto ADE, è stata promossa da John Tukey nel 1961 definendola come quell'insieme di: *procedure per analizzare i dati, di tecniche per interpretare i risultati di tali procedure, di metodi per pianificare la raccolta dei dati al fine di rendere l'analisi più facile, più precisa e più accurata, e di tutti gli strumenti e risultati delle statistiche (matematiche) che si applicano per l'analisi dei dati.*¹⁴ Nei suoi lavori, egli sottolinea l'importanza di questo approccio e lo differenzia dall'analisi inferenziale, ossia la verifica delle ipotesi statistiche, detta anche analisi dei dati confermativa (ADC), secondo l'autore, spesso sopravvalutata. Mentre l'analisi inferenziale segue la definizione dell'ipotesi di ricerca, l'analisi esplorativa normalmente ha luogo prima della definizione del modello teorico e quindi dell'ipotesi di ricerca stessa. In termini procedurali, questo significa che nella ADC, la sequenza è rappresentata da: *Problema -> definizione di ipotesi sulla base di un modello teorico -> raccolta dati -> analisi -> conclusioni (eventuali)*; nell'ADE si ha: *Problema -> raccolta dati -> analisi esplorativa -> definizione di un modello (ipotesi) -> (eventuali) conclusioni.*

In letteratura, è facile ritrovare il termine analisi descrittiva/esplorativa, ma una lettura più attenta si evidenzia una distinzione tra analisi descrittiva, ossia l'uso di tecniche di analisi univariata non inferenziale, volte cioè allo studio dell'impatto di

¹⁴ Tukey, John W. The Future of Data Analysis. Ann. Math. Statist. 33 (1962), no. 1, 1--67

una sola variabile sul fenomeno studiato, e analisi esplorativa, e cioè l'uso di tecniche multivariate per valutare la relazione fra due o più variabili, per suggerire ipotesi da testare. Secondo Tukey, gli specifici obiettivi della ADE sono:

- Suggerire ipotesi sulle cause dei fenomeni osservati
- Valutare presupposti su cui si baserà l'inferenza statistica
- Sostenere la selezione di adeguati strumenti e tecniche statistiche
- Fornire una base per ulteriore raccolta di dati attraverso indagini o esperimenti

In sintesi quindi, in questa fase, sono state utilizzate tecniche di statistica esplorativa multivariata al fine di *individuare relazioni tra i dati non banali, nascoste, utili e fruibili dall'utilizzatore* (David Hand (2000)). In altre parole, con questa analisi si è andati alla ricerca di aspetti non osservabili direttamente con gli strumenti della statistica classica e che permettono, tra le tante, il trattamento simultaneo di numerose variabili, la visualizzazione di associazioni complesse, l'individuazione di fattori multidimensionali, la definizione delle priorità da affrontare e studiare.

Questa fase della ricerca è stata articolata in 4 momenti, in ognuno dei quali si è fatto ricorso a una o più tecniche specifiche finalizzate ad analizzare una criticità, e tutte utilizzate senza un'ipotesi di ricerca, e quindi non in termini inferenziali. Di seguito si riporta una breve descrizione delle **tecniche utilizzate**:

1) Chi Quadro: in statistica, test o indice di adattamento (indicato con il simbolo χ^2) introdotto da K. Pearson allo scopo di valutare il grado di adattamento di una certa distribuzione teorica ad una osservata. In altre parole permette di misurare se le differenze fra una distribuzione effettiva, osservata, e la corrispondente distribuzione teorica, ossia la distribuzione ottenuta idealmente sulla base dell'ipotesi teorica, siano dovute al caso oppure siano significative. Può essere utilizzata sia per verificare in che misura una distribuzione campionaria è coerente con quella teorica di riferimento, così da decidere ad esempio, se un campione fa parte di una data popolazione con quella distribuzione, oppure può essere impiegato per verificare l'attendibilità di una relazione tra due fenomeni, cioè per verificare, ad esempio, se i due fenomeni siano o non siano significativamente connessi. Il χ^2 si calcola con la formula:

Indice di connessione del Pearson

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - n_{ij}^e)^2}{n_{ij}^e}$$

$$0 \leq \chi^2 \leq N \times \underbrace{[\min(k, h) - 1]}_{\text{Il più piccolo valore tra il numero di righe e il numero di colonne della tabella meno 1}}$$

e si basa sulla tabella di contingenza, calcolata come: $(f_o - f_a)^2 / f_a$ ¹⁵. “Il termine *tabella di contingenza* (in inglese *contingency table*) deriva dal fatto che ogni distribuzione, in riga o in colonna, è limitata ai valori su una variabile dei soli dati che presentano un certo valore sull’altra variabile. Si tratta cioè di distribuzioni di frequenza condizionate (in inglese *contingent*)”. (Marradi, 1997). Questo test ci consente di stabilire (dopo aver fissato l’errore massimo tollerato) se le discrepanze tra le frequenze osservate e quelle teoriche sono imputabili completamente al caso o se invece è lecito supporre che ci sia una qualche forma di dipendenza tra i dati. Se $\chi^2 = 0$, le frequenze osservate coincidono esattamente con quelle teoriche. Se invece $\chi^2 > 0$, esse differiscono. Più grande è il valore di χ^2 , più grande è la discrepanza tra le frequenze osservate e quelle teoriche. È evidente che il chi-quadrato aumenta con l’aumentare della differenza dei dati posti a raffronto. Se esso supera certi valori prefissati (vedi tabella «Valori di chi-quadrato»), la differenza viene ritenuta significativa; in caso contrario, non si può affermare l’esistenza di una significativa differenza.

2) V di Cramer: proposto dal matematico Gabriel Cramer, si tratta di un indice simmetrico di associazione, ottenuto come radice quadrata del χ^2 , rapportato alla frequenza totale e al valore minimo tra $r-1$ e $c-1$, dove r e c sono il numero di modalità distinte del primo e, rispettivamente, del secondo carattere, e fornisce indicazioni sul livello di associazione non strutturata tra caratteri, particolarmente qualitativi nominali. La formula è mostrata in Figura 11. Come il chi di quadro, il V di Cramer misura la connessione tra variabili qualitative, ma diversamente dal chi quadro, esso assume valori finiti che vanno da 0 (nessuna connessione) a 1 (massima connessione).

¹⁵ Tabella di Contingenza = $(f_o - f_a)^2 / f_a$; dove f_o sta per frequenze osservate ed f_a per frequenze attese o teoriche.

<p>Indice di contingenza media quadratica</p> $\phi^2 = \frac{\chi^2}{N} = \frac{1}{N} \sum_i \sum_j \frac{(n_{ij} - n_{i.} n_{.j})^2}{n_{ij}}$ $0 \leq \phi^2 \leq \min(k, h) - 1$ <p>Indice V di Cramer</p> $V = \sqrt{\frac{\phi^2}{\min(h, k) - 1}}$ $0 \leq V \leq 1$
--

Figura 11: Formula della V di Cramer

3) Regressione logistica: caso particolare di modello lineare generalizzato ossia un modello di regressione applicato quando la variabile dipendente y è di tipo dicotomico, riconducibile ai valori 0 e 1, e che quindi può assumere solo 2 valori, come ad esempio vero o falso, si o no, maschio o femmina, ecc.. La regressione logistica ci permette di misurare in che percentuale all'aumentare di una data variabile, X , una seconda variabile Y , dicotomica, ad essa connessa, varia.

4) Analisi delle Corrispondenze (AC): è una delle tecniche dell'analisi fattoriale. L'Analisi fattoriale (AF) permette di ottenere una riduzione della complessità del numero di fattori che spiegano un fenomeno. In altre parole, si propone di estrarre un numero di variabili "latenti" inferiore e riassuntivo rispetto al numero di variabili di partenza, basandosi sulla parte comune delle rilevazioni fatte, e consente inoltre di costruire dei modelli teorico-matematici capaci di fornire molte informazioni circa i rapporti di causalità tra le variabili (Giannini e Pannocchia, 2006; Lubisco, 2006, 2007). Nella sua accezione esplorativa, l'AFE è largamente utilizzata per formulare ipotesi circa le dimensioni latenti sottostanti ai costrutti sui quali sono costruiti i test. Esistono diversi tipi di analisi fattoriale, in base ai dati utilizzati in ingresso. Una di queste è l'analisi delle corrispondenze, il scopo specifico è di *individuare dimensioni soggiacenti alla struttura dei dati, ossia dimensioni in grado di riassumere l'intreccio di relazioni di "interdipendenza" tra le variabili originarie. Tramite l'analisi delle corrispondenze si può trasformare una tabella di contingenza in una rappresentazione grafica al fine di facilitare l'interpretazione delle informazioni contenute nella tabella stessa*¹⁶. Introdotta negli anni '70, ad opera della scuola francese di analisi dei dati (Benzecri), fu proposta per analizzare tabelle di

¹⁶ Da Wikipedia

contingenza bivariate di grandi dimensioni. Questa tecnica ha assunto una grande importanza in psicologia, in quanto può essere utilizzata sia con dati qualitativi che quantitativi, che come si sa caratterizzato gli strumenti di valutazione psicologica, tipo i questionari. Successivamente è stata sviluppata una versione più complessa, l'analisi delle corrispondenze multivariate (ACM) per tabelle contenenti informazioni relative ad un numero elevato di variabili.

5) Analisi delle Componenti Principali (ACP): L'ACP è una delle tecniche dell'analisi fattoriale, (vedi sopra), proposta nel 1901 da Karl Pearson e sviluppata da Harold Hotelling nel 1933, e come tale finalizzata alla semplificazione dei dati nell'ambito della statistica multivariata. Lo scopo primario di questa tecnica è ridurre un certo numero, più o meno elevato, di variabili (rappresentanti altrettante caratteristiche del fenomeno analizzato) in alcune variabili latenti (feature reduction). Questa riduzione si ottiene mediante una trasformazione lineare delle variabili che proietta quelle originarie in un nuovo sistema cartesiano nel quale la nuova variabile con la maggiore varianza viene proiettata sul primo asse, la seconda sul secondo asse e così via, in base alla dimensione della varianza. La riduzione della complessità avviene limitandosi ad analizzare le principali (per varianza) tra le nuove variabili. In altre parole, le variabili osservate sono sostituite dalle nuove componenti principali, che insita nella loro costruzione hanno come vantaggi i) di non essere correlate tra loro; ii) di essere presentate in ordine decrescente di varianza e iii) che la varianza totale (somma delle varianze) si conserva nel passaggio dalle variabili osservate alle componenti principali. Proprio per queste proprietà si può affermare che le componenti principali forniscono una spiegazione alternativa della variabilità osservata con il pregio di descrivere il fenomeno attraverso dimensioni fra loro non correlate e ordinate in termini della loro importanza nella spiegazione. Questo consente di diminuire il numero di variabili da considerare, scartando quelle il cui contributo è trascurabile alla spiegazione della variabilità osservata, ed inoltre di interpretare il fenomeno attraverso un'interpretazione delle principali. Il numero di CP da considerare varia da caso a caso in base sia alla percentuale di varianza totale spiegata da ogni componente principale, e sia in relazione a quella spiegata dalle altre, oltre che alla conoscenza specifica del fenomeno oggetto di studio e all'uso che

se ne vuole fare. Spesso in questa decisione aiuta lo screen plot, che mostra graficamente le differenze significative.

6) Analisi delle Correlazioni Canoniche (ACC): proposta da H. Hotelling nel 1936 allo scopo di analizzare i legami tra le caratteristiche fisiologiche (variabili antropometriche) e quelle mentali (risposte ai test psicologici) di un insieme di individui, è stata poi ampliata e generalizzata da due a tre o più gruppi di variabili da J. D. Carroll, nel 1968. L'obiettivo dell'ACC è di identificare le relazioni lineari esistenti tra due gruppi di variabili osservate su uno stesso insieme di individui. Lo scopo è trovare una combinazione lineare delle variabili del primo gruppo e una combinazione lineare delle variabili del secondo che siano massimamente correlate, definite coppie di componenti canoniche. Il metodo analizza cioè le interdipendenze tra due gruppi di variabili e tenta di individuare un doppio sistema di variabili latenti che riproducano la correlazione tra i due gruppi di variabili osservate, senza quella presente al loro interno. In altre parole, la rappresentazione originaria dei due gruppi di variabili misurate, viene sintetizzata e trasformata in una nuova rappresentazione data dalla creazione di due nuovi gruppi di variabili artificiali, che non presentano correlazioni al loro interno ma sono massimamente correlate tra loro. Le nuove variabili canoniche possono, a questo punto, essere usate per ulteriori analisi, o rappresentate graficamente. Sulla base del primo autovalore estratto, e cioè della correlazione canonica più alta, è possibile effettuare un test di indipendenza fra i due sottoinsiemi di variabili.

Per alcune di queste tecniche non sono state fornite le formule per il calcolo, perché la presente ricerca si focalizza sugli aspetti interpretativi forniti da queste, piuttosto che sulle loro caratteristiche statistiche. Inoltre, essendo molto complesse, sono calcolate normalmente tramite strumenti statistici specifici, quali SPSS, e soprattutto R. Di seguito la descrizione del loro uso nelle diverse fasi dell'analisi.

Per quanto riguarda la base dati utilizzata, il **campione dei dati**, si fa notare che a causa di una differenza nello strumento, descritta nel paragrafo 3.2.5, in questa sede sono presi in esame solo i risultati del secondo trial, in quanto più conformi alla teoria e alle tecniche di Rahim.

3.3.2 Preparazione dei data set

Il primo passo propedeutico all'analisi è stata la riorganizzazione dei dati in una serie di data set utili. Questo passaggio è risultato necessario, in quanto, come si è già detto, ENACT presenta un database molto articolato e complesso, con dati raccolti in maniera automatica e randomica (vedi Figura 12), ed è stato pertanto necessario trattare in maniera diversa a seconda delle analisi da effettuare. Il dataset iniziale, ossia i dati così come sono stati raccolti dal sistema era così **composto**:

1. Anagrafica: nome partner che ha condotto la sperimentazione, username dell'utente, ID, genere, lingua, età, nazione, minor età;
2. Dati ENACT: ordine randomico della presentazione degli scenari, per ciascuno scenario, tre tipi di informazione per ciascun item: Type (stile della risposta), Pos (posizione dell'item nel pattern di 5 proposto); Time (decimi di secondo impiegati per la scelta);
3. Risultati ROCI II – Form C: 28 item per ciascuno una risposta compresa tra 1 (per niente d'accordo) e 5 (completamente d'accordo).

	I	J	K	L	M	N	O	P	Q	R	S	CM	CN	CO	CP	CC
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	1-1 type	1-1 pos	1-1 time	1-2 type	1-2 pos	1-2 time	ROCI 1	ROCI 2	ROCI 3	ROCI 4	ROCI 5
2	3	1	2	4	5	1	1	1135	1	3	1181	3	3	5	3	3
3	2	4	5	1	3	1	2	734	4	2	351	5	5	5	5	1
4	3	4	5	2	1	2	3	5502	4	4	2430	4	1	1	4	4
5	3	1	4	5	2	1	1	2148	1	0	783	4	4	3	4	3
6	3	2	1	5	4	1	0	4510	5	4	1304	5	3	5	4	3
7	4	2	1	5	1	3	1	704	3	0	1347	5	5	3	4	4
8	2	3	4	1	5	5	4	2116	5	2	1527	4	3	1	4	4
9	2	5	1	4	3	5	2	2090	5	2	2724	4	1	1	4	5
10	1	2	3	5	4	1	4	2536	1	3	2945	5	3	2	4	4
11	3	5	4	2	1	1	0	2579	5	1	2123	5	5	4	5	3
12	4	5	3	2	1	1	0	790	1	0	2374	4	4	4	3	4

Figura 12: Differenza tra i due set di dati

Come si evince chiaramente dalla figura, a differenza del ROCI i dati di ENACT, riportano un numero maggiore di informazioni, di natura diversa. Per effettuare le analisi i dati sono stati riorganizzati in 3 database diversi a seconda delle esigenze di analisi:

- 1) **Dati per stile**: i dati, sia quelli di ENACT che del ROCI, sono stati riorganizzati per stile (vedi Figura 13) ossia è stata calcolata per ciascun soggetto la frequenza di scelta di ciascuno stile nei diversi scenari. Ricordiamo che mentre per il ROCI si tratta di punteggi misurati su scala Likert, nel caso di ENACT si tratta di frequenze di risposta.

Roci										ENACT																																										
Integrating		Dominating		Obliging		Avoiding		Compromising		Integrating		Dominating		Obliging		Avoiding		Compromising																																		
1	4	5	12	22	23	28	8	9	18	21	25	2	10	11	13	19	24	3	6	16	17	26	27	7	14	15	20	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
3	3	3	3	4	3	3	3	1	2	4	2	3	3	4	3	3	3	5	5	5	4	4	5	3	3	3	3	3	2	4	1	3	1	2	0	1	1	0	0	0	3	0	0	0	0	1	1	1	0	0		
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	1	5	1	1	1	5	5	5	5	0	2	1	0	0	0	2	0	0	0	0	0	1	0	2	3	0	3	3	3	2	1	0	2	0
4	4	4	3	4	4	4	4	4	3	3	3	3	4	3	2	4	4	3	2	4	2	3	2	4	4	4	5	0	3	0	0	0	0	0	2	0	0	2	2	0	1	2	2	0	3	1	2	1	0	0	3	1
4	4	3	4	4	4	3	3	3	4	3	5	4	3	1	2	3	4	3	3	5	1	5	3	5	4	4	4	4	3	3	0	0	0	1	2	0	0	0	0	1	0	1	2	2	1	1	1	0	2	1		
5	4	5	5	5	5	4	1	1	3	4	5	3	1	2	3	5	4	5	5	5	1	3	2	4	3	4	4	4	4	2	0	0	0	1	0	0	0	1	0	0	1	0	1	1	4	1	0	2	1	0		
5	4	4	4	4	5	5	3	2	3	2	3	5	3	2	3	4	5	3	3	5	3	2	5	4	5	3	4	0	3	2	2	0	0	0	0	0	5	2	2	0	5	0	0	0	1	0	0	1	2	0		
4	4	4	4	4	4	4	2	1	1	3	1	3	1	2	2	3	1	1	2	3	3	1	3	1	2	1	1	1	3	3	1	0	0	1	0	0	3	4	0	0	4	1	0	0	0	1	0	1	2	0	0	
5	4	5	5	5	5	5	1	1	2	2	4	3	3	3	2	4	4	3	3	3	3	4	5	2	5	4	3	1	1	3	2	2	0	0	0	0	0	3	2	0	1	1	0	0	0	2	2	1	2	2	0	0
5	4	4	5	5	5	4	3	1	4	1	3	3	2	2	3	4	4	2	3	5	4	2	2	3	4	3	5	4	2	4	0	1	0	0	0	0	0	1	0	4	0	0	0	2	1	2	1	1	2			
5	5	5	4	5	5	5	4	5	5	1	5	5	4	5	4	5	4	5	5	5	5	5	5	5	4	5	5	0	3	1	1	0	0	0	1	0	0	4	0	0	2	4	1	1	0	1	1	0	1	3	1	0

Scenario		
N scenario	Titolo	Stile BOT
1	Sport	OBL
2	Motorbike	INT
3	Restaurant	COM
4	CD	AVO
5	TVProgramme	DOM

Figura 13: Dati per stile

Tutte le misure del test e il confronto con il ROCI effettuate nel progetto e riportate nei deliverable, ad es. corrispondenze, Alpha di Cronbach, validità, ecc. (vedi paragrafo 3.2.5) sono state calcolate partendo da questo database, che permette di confrontare i punteggi ottenuti dai soggetti sui diversi stili. Ricordiamo che gli assunti che stanno alla base di questa scelta, sono che 1) il ROCI non ha scenari, ma tiene conto solo degli stili, per cui 2) lo scenario può essere visto come un item a più risposte, e che, pertanto, 3) la frequenza di scelta di un certo stile in un determinato scenario può rappresentare un punteggio quantitativo del soggetto per quello stile in quello scenario. Un'analisi dettagliata di questo assunto e delle sue implicazioni è riportata nel paragrafo 3.2.6.1.

2) **Dati per Scenario:** i dati di ENACT sono stati riorganizzati per tipo di scenario. Naturalmente questo non può essere fatto per il ROCI che non prevede questa modalità. In altre parole a ciascuno stile è stato dato un codice numerico, e riorganizzati in base allo scenario ordinandoli per numero di risposta. In questo, e più in generale in tutti i database, sono stati esclusi i dati relativi alla posizione dell'item nel pattern e al tempo di risposta sia perché l'analisi dei dati originali ha dimostrato che queste variabili non avevano effetti significativi, sia per semplificare il dataset, già complicato dalla presenza degli scenari, sia per effettuare un'analisi critica solo sui risultati relativi agli stili, parte centrale e fondamentale del test, oltre che unica variabile utili per il confronto con il test originale.

Ordinati per tipo di scenario																							
SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE	SCE
1_ri	1_ri	1_ri	1_ri	1_ri	2_ri	2_ri	2_ri	2_ri	2_ri	3_ri	3_ri	3_ri	3_ri	3_ri	4_ri	4_ri	4_ri	4_ri	4_ri	5_ri	5_ri	5_ri	5_ri
sp1	sp2	sp3	sp4	sp5	sp1	sp2	sp3	sp4	sp5	sp1	sp2	sp3	sp4	sp5	sp1	sp2	sp3	sp4	sp5	sp1	sp2	sp3	sp4
2	1	5	1	1	2	1	2	1	5	1	1	5	1	1	1	3	3	2	3	4	2	1	1
5	4	5	4	4	1	2	5	1	2	1	4	4	4	3	5	4	5	4	4	4	4	3	4
4	5	4	3	3	3	1	3	1	1	2	4	4	4	2	3	4	5	5	5	4	5	4	3
5	1	1	1	1	4	5	1	1	1	1	1	1	4	4	2	4	5	5	4	4	2	3	5
5	1	1	1	1	4	1	1	1	1	1	5	5	4	1	3	5	4	1	1	4	4	4	4
3	3	3	3	3	1	3	1	3	1	3	3	1	1	5	1	1	4	5	5	3	3	3	3
1	3	3	3	3	5	2	1	1	1	5	5	1	1	1	1	3	3	3	3	4	2	3	2
3	3	5	1	3	5	3	5	1	3	5	5	1	1	1	1	4	4	1	3	4	3	1	4
1	1	1	5	1	1	5	3	1	5	1	1	1	5	1	3	3	5	3	3	4	4	5	5
3	4	3	3	3	1	1	1	4	5	1	5	2	5	5	1	5	4	3	3	3	4	3	3

stile risposta	
N risposta	nome stile
1	INT
2	DOM
3	OBL
4	AVO
5	COM

Figura 14: Dati ENACT per Scenario

La difficoltà di questo raggruppamento nasce dal fatto che nel corso della sperimentazione gli scenari sono presentati in maniera random e quindi nella raccolta, come si vede dalla Figura 12, per ciascun soggetto viene riportato il numero di scenario e i punteggi ottenuti in base alla sequenza di esecuzione. Questa sequenza è diversa per ciascun soggetto. Nel nuovo database (Figura 14), è stato necessario eliminare l'effetto randomico e aggregare i dati in base al numero di scenario.

3) **Dati con etichette:** per svolgere alcuni tipi di analisi qualitative è stato necessario etichettare i valori ottenuti su ENACT in base allo stile di risposta, piuttosto che utilizzare il codice numerico utilizzati nei precedenti dataset (Figura 15).

ordinati per tipo di scenario con inserimento etichette stile																									
scenario 1 - SPORT (B=Obl)					scenario 2 - MOTORBIKE					scenario 3 - RESTOURANT					scenario 4 - CD (B=Avo)					scenario 5 - TV PROGRAMME					
R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	R1	R2	R3	R4	R5	
Dom	Int	Com	Int	Int	Dom	Int	Dom	Int	Com	Int	Int	Com	Int	Int	Int	Obl	Obl	Dom	Obl	Avo	Dom	Int	Int	Int	
Com	Avo	Com	Avo	Avo	Int	Dom	Com	Int	Dom	Int	Avo	Avo	Avo	Obl	Com	Avo	Com	Avo	Avo	Avo	Avo	Obl	Avo	Obl	
Avo	Com	Avo	Obl	Obl	Obl	Int	Obl	Int	Int	Dom	Avo	Avo	Avo	Dom	Obl	Avo	Com	Com	Com	Avo	Com	Avo	Obl	Obl	
Com	Int	Int	Int	Int	Avo	Com	Int	Int	Int	Int	Int	Int	Avo	Avo	Dom	Avo	Com	Com	Avo	Avo	Dom	Obl	Com	Dom	
Com	Int	Int	Int	Int	Avo	Int	Int	Int	Int	Int	Com	Com	Avo	Int	Obl	Com	Avo	Int	Int	Avo	Avo	Avo	Avo	Dom	
Obl	Obl	Obl	Obl	Obl	Int	Obl	Int	Obl	Int	Obl	Obl	Int	Int	Com	Int	Int	Avo	Com	Com	Com	Obl	Obl	Obl	Obl	
Int	Obl	Obl	Obl	Obl	Com	Dom	Int	Int	Int	Com	Com	Int	Int	Int	Int	Obl	Obl	Obl	Obl	Obl	Avo	Dom	Obl	Dom	Dom
Obl	Obl	Com	Int	Obl	Com	Obl	Com	Int	Obl	Com	Com	Int	Int	Int	Int	Avo	Avo	Int	Obl	Avo	Obl	Int	Avo	Int	
Int	Int	Int	Com	Int	Int	Com	Obl	Int	Com	Int	Int	Int	Com	Int	Obl	Obl	Com	Obl	Obl	Avo	Avo	Com	Com	Int	
Obl	Avo	Obl	Obl	Obl	Int	Int	Int	Avo	Com	Int	Com	Dom	Com	Com	Int	Com	Avo	Obl	Obl	Obl	Avo	Obl	Obl	Obl	
Com	Obl	Avo	Com	Com	Int	Int	Obl	Int	Int	Int	Int	Int	Int	Com	Int	Int	Com	Obl	Obl	Com	Obl	Avo	Com	Com	
Int	Com	Int	Int	Int	Avo	Int	Obl	Int	Int	Obl	Int	Obl	Int	Obl	Int	Obl	Obl	Com	Obl	Avo	Obl	Dom	Int	Com	
Com	Com	Int	Int	Int	Int	Obl	Com	Int	Int	Int	Com	Com	Int	Int	Int	Int	Obl	Com	Int	Int	Avo	Com	Obl	Int	

Figura 15: Dati ENACT con Etichette Stile

Di seguito sono descritte le analisi svolte sui diversi dataset al fine di rispondere ai quesiti sorti in seguito all'analisi critica dei risultati del progetto.

3.3.3 Analisi 1

“Studio dell’indipendenza delle risposte mediante le statistiche del Chi Quadro e V di Cramer”.

D: I punteggi ottenuti dai soggetti nei singoli scenari di ENACT possono essere considerati misure indipendenti oppure data la struttura del test, i punteggi devono tener conto della consequenzialità delle risposte e quindi devono essere considerate dipendenti tra loro?

Come abbiamo già avuto modo di descrivere nei paragrafi precedenti (3.2.6.1), la prima differenza importante tra i due strumenti riguarda l’indipendenza dei dati. I punteggi del ROCI, secondo la teoria di Rahim, sono indipendenti tra loro, e le medie dei punteggi ottenuti su ciascun gruppo di item specifico per ciascuno stile, danno la misura quantitativa dello stile di gestione del conflitto da parte di quel dato soggetto. Data la diversa struttura dei due strumenti, ENACT propone scenari, con domande interconnesse tra loro, ma possiamo affermare che questo sia vero anche per ENACT?

La prima analisi si è quindi focalizzata ad identificare quanto gli item di ENACT all’interno dello stesso scenario sono dipendenti tra loro.

Il principio utilizzato per le analisi effettuate nel progetto è che ogni risposta, delle 5 possibili, rappresenta un punteggio sulla scala che misura quello specifico stile, per cui la somma delle risposte dello stesso stile per ciascuno scenario può essere considerato come una misura quantitativa. Per esemplificare si fa riferimento alla Figura 13: Dati per stile che mostra i risultati ottenuti dai soggetti organizzati per stile di risposta, nei diversi scenari.

In altre parole, si è assunto che se un soggetto ha scelto lo stile Integrating 4 volte nello scenario 1, 3 nello scenario 2, 2 volte nel 3, 1 volta nel 4 e 3 volte nel 5, questi valori possono essere considerati i punteggi ottenuti ad un gruppo di item che misurano lo stile Integrating nell’intero test, considerandola alla stregua di una scala Likert a 5 passi. La scala Likert fu ideata dallo psicometrico americano Rensis Likert nel 1932 con lo scopo di elaborare *un nuovo strumento, più semplice rispetto ad altri, per la misurazione di opinioni e atteggiamenti* (Likert, 1932). Nel caso specifico, la frequenza di risposta è vista come il punteggio ottenuto. Le

considerazioni alla base di questo metodo sono che: i) il ROCI non ha scenari, ma tiene conto solo degli stili, per questo ii) il singolo scenario è considerato un item “a più risposte” e che iii) più volte uno stile viene scelto dal soggetto in uno specifico scenario, tanto più è alto il punteggio ottenuto dal soggetto per quello stile su una scala di misura simile a quella Likert utilizzata per il ROCI.

Questo metodo è fortemente opinabile. Innanzitutto, la frequenza dello stile di risposta non può essere considerato come un punteggio definito, come lo è ad esempio un 4 ottenuto ad un item di Rahim, e quindi le due misure non sono immediatamente confrontabili. I punteggi del ROCI infatti, seguendo la teoria classica, possono essere considerate misure quantitative, mentre la frequenza è una misura qualitativa. Inoltre, anche se si volesse ritenere che la frequenza, in questo caso specifico, possa essere considerata una misura quantitativa, partendo dagli assunti sopra descritti, è necessario sottolineare che un punteggio 4 ottenuto con il ROCI è una misura unica e indipendente, in quanto risposta ad un solo item, mentre in ENACT è la somma di 4 risposte date dal soggetto all'interno dello stesso scenario. Poiché in via teorica, si assume che ci sia una certa dipendenza tra le risposte consecutive date ad uno stesso scenario, è necessario verificarne l'esistenza. Questo perché se si dimostra una forte dipendenza tra le risposte questa implica che:

- 1) è necessario tenere conto di questa condizione, e trovare un metodo che tenga conto di questa dipendenza nel calcolo dei punteggi finali
- 2) tutte le analisi precedenti e successive devono essere ricalcolate sulla base di questi nuovi valori.

In vista di queste considerazioni, per misurare l'indipendenza dei dati di ENACT per ciascuno scenario si è fatto ricorso alle statistiche specifiche per le frequenze, in particolare il Chi Quadro. L'obiettivo finale è verificare quanto le risposte sono interconnesse tra loro.

Utilizzando il dataset organizzato per scenario (Figura 14), si è proceduto a calcolare le frequenze e le contingenze, per ciascun scambio per ciascuno scenario, al fine di identificare e quantificare le dipendenze tra le risposte a ciascuno scenario.

È importante notare che in questa analisi è stata esclusa l'ultima interazione, tra risposta 4 e 5, poiché, l'item 5 è di natura diversa e non può essere considerato alla

stessa stregua degli altri. Una descrizione dettagliata è fornita nel paragrafo 3.2.6.1. Questo porta a considerare solo le prime tre interazioni.

Per ciascun scenario, sono stati incrociati i dati a coppie successive di risposte, ottenendo così per ciascuna coppia una tabella di frequenze osservate, per ognuna delle quali sono state calcolate le frequenze attese e successivamente le contingenze (vedi Figura 16). Le contingenze sono alla base del calcolo del CHI Quadro (2), e permettono quindi di definire in che misura le co-occorrenze che si verificano sono significative o semplicemente legate al caso.

È bene notare che il test del 2 non si può applicare se più del 20% delle celle ha una “frequenza attesa” inferiore a 5¹⁷; oppure se una cella ha frequenza attesa inferiore a 1. In questo caso, se vi sono 3 o più categorie e se il significato logico di tali categorie lo permette, è possibile aggregarne alcune in modo da ampliare la numerosità di quella particolare riga/colonna.

Scenario 1 - SPORT FREQUENZE OSSERVATE																								
		Risp 2								Risp 3								Risp 4						
Risp 1	INT	DOM	OBL	AVO	COM	TOT			Risp 2	INT	DOM	OBL	AVO	COM	TOT			Risp 3	INT	DOM	OBL	AVO	COM	TOT
INT	28	5	9	6	8	56			INT	28	1	5	4	10	48			INT	40	1	10	1	3	55
DOM	3	2	2	1	2	10			DOM	3	3	2	1	2	11			DOM	3	4	1	0	1	9
OBL	7	3	10	3	3	26			OBL	9	1	13	4	4	31			OBL	6	0	14	5	3	28
AVO	2	0	6	1	8	17			AVO	6	3	2	1	3	15			AVO	1	0	6	3	5	15
COM	8	1	4	4	3	20			COM	9	1	6	5	3	24			COM	11	3	2	2	4	22
TOT	48	11	31	15	24	129			TOT	55	9	28	15	22	129			TOT	61	8	33	11	16	129

Scenario 1 - SPORT FREQUENZE ATTESE																								
		Risp 2								Risp 3								Risp 4						
Risp 1	INT	DOM	OBL	AVO	COM	TOT			Risp 2	INT	DOM	OBL	AVO	COM	TOT			Risp 3	INT	DOM	OBL	AVO	COM	TOT
INT	21	5	13	7	10	56			INT	20	3	10	6	8	48			INT	26	3	14	5	7	55
DOM	4	1	2	1	2	10			DOM	5	1	2	1	2	11			DOM	4	1	2	1	1	9
OBL	10	2	6	3	5	26			OBL	13	2	7	4	5	31			OBL	13	2	7	2	3	28
AVO	6	1	4	2	3	17			AVO	6	1	3	2	3	15			AVO	7	1	4	1	2	15
COM	7	2	5	2	4	20			COM	10	2	5	3	4	24			COM	10	1	6	2	3	22
TOT	48	11	31	15	24	129			TOT	55	11	31	15	24	129			TOT	61	11	31	15	24	129

Scenario 1 - SPORT Tavola delle Contingenze																								
		Risp 2								Risp 3								Risp 4						
Risp 1	INT	DOM	OBL	AVO	COM				Risp 2	INT	DOM	OBL	AVO	COM				Risp 3	INT	DOM	OBL	AVO	COM	
INT	2,46	0,01	1,48	0,04	0,56				INT	2,77	1,65	2,82	0,45	0,40				INT	7,53	1,70	1,18	2,90	2,14	
DOM	0,14	1,54	0,07	0,02	0,01				DOM	0,61	6,49	0,06	0,06	0,01				DOM	0,37	21,22	0,74	0,77	0,01	
OBL	0,74	0,28	2,25	0,00	0,70				OBL	1,35	0,63	5,85	0,04	0,31				OBL	3,96	1,74	6,53	2,86	0,06	
AVO	2,96	1,45	0,90	0,48	7,40				AVO	0,02	3,65	0,48	0,32	0,08				AVO	5,23	0,93	1,22	2,32	5,30	
COM	0,04	0,29	0,14	1,21	0,14				COM	0,15	0,27	0,12	1,75	0,29				COM	0,03	1,96	2,34	0,01	0,59	

Figura 16: Frequenze osservate, attese e contingenze – Scenario 1- ENACT

Se si osservano i nostri dati, (vedi Tabella 7) si può notare che in tutte le tabelle i) c'è un numero superiore al 20% di celle con valori inferiori al 5 e ii) che nella

¹⁷solo nel caso di una tabella 2x2, si può utilizzare una formula alternativa chiamata “correzione di Yates”

maggioranza dei casi, ci sono celle con valori inferiori a 1. Anche raggruppando i valori delle prime due risposte la situazione non cambia. Questo ci induce ad affermare che, data la natura dei nostri dati, non è possibile utilizzare il test del χ^2 ai fini inferenziali ma solo come una misura descrittiva.

	Scenario 1				Scenario 2				Scenario 3				Scenario 4				Scenario 5			
	2	<5	%	<1	2	<5	%	<1	2	<5	%	<1	2	<5	%	<1	2	<5	%	<1
Int 1	25,30	17	68%	1	18,47	17	68%	3	39,92	18	72%	5	17,56	15	60%	2	26,23	16	64%	1
Int 2	30,63	15	60%	1	25,13	18	72%	12	51,28	17	68%	3	32,08	12	48%	1	26,16	15	60%	0
Int 3	73,64	17	68%	3	24,34	21	84%	12	50,98	17	68%	2	44,40	12	48%	0	47,70	15	60%	0

Tabella 7: Chi quadro e frequenze attese nei 5 scenari di ENACT

Tuttavia, dato che il χ^2 non ha un massimo definito, ed è quindi di più difficile interpretazione, si è deciso di ricorrere ad un altro indice descrittivo di connessione, il V di Cramer, che consiste in una versione normalizzata del χ^2 ed quindi pari a 0 nel caso di indipendenza e pari a 1 nel caso di massima connessione.

Come si evince dalla Tabella 8, i valori del V di Cramer, nelle diverse interazioni, vanno da un minimo di 0.18 ad un massimo di 0.38, con una media che non supera mai lo 0,30. Questo ci porta ad affermare che non c'è una dipendenza significativa tra le coppie di risposte consecutive nei singoli scenari.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Media
Int 1	0,22	0,19	0,28	0,18	0,23	0,22
Int 2	0,24	0,22	0,32	0,25	0,23	0,25
Int 3	0,38	0,22	0,31	0,29	0,30	0,30
Media	0,28	0,21	0,30	0,24	0,25	

Tabella 8: V di Cramer negli scenari di ENACT

Conclusione analisi 1: La conclusione di questo lavoro è duplice: da una parte, l'impossibilità di utilizzare il chi quadro per fini inferenziali, ci induce ad affermare che sarebbe necessario **aumentare il numero dei partecipanti al test, al fine di ottenere un campione maggiormente rappresentativo.** Dall'altro, il calcolo del V di Cramer, come indice di connessione normalizzato, ci porta a dedurre che non c'è una forte dipendenza tra le coppie di risposte, di conseguenza gli item appaiono indipendenti l'uno dall'altro ed è pertanto **possibile utilizzare nelle analisi i punteggi originali.** In aggiunta, si può quindi affermare che **tutte le statistiche calcolate in prima istanza (a fine progetto), hanno pertanto una validità statistica.**

3.3.4 Analisi 2

“Studio della validità dell’item di controllo con la Regressione logistica”

D: I quattro item di ciascuno scenario sono coerenti con l’item di controllo?

Come abbiamo già evidenziato, un aspetto importante di ENACT è rappresentato dalla differenza degli item che compongono i singoli scenari. Delle 5 domande consecutive che sono poste ai soggetti in ogni scenario, la quinta rappresenta una sintesi del comportamento, una misura di controllo della coerenza dello scenario, una sorta di item di controllo. Quindi, sebbene consecutiva e logicamente connessa alle precedenti, ha una natura completamente diversa.

Per una descrizione dettagliata di questo item si rimanda al paragrafo 3.2.6.1, in cui si mostra una tipica interazione dello scenario, “programmi TV”, in cui si evince la diversa caratterizzazione dei 5 item. In questa sezione, invece si analizza l’impatto che l’item n. 5 ha sulle analisi dei dati, a causa delle sue peculiarità, e in particolare si indaga sulla coerenza interna degli scenari.

A tal fine, per verificare se l’item n. 5 è stato costruito in maniera coerente rispetto agli altri 4 presenti nello scenario, si è deciso di ricorrere alla regressione logistica. Nel nostro caso specifico, abbiamo effettuato una serie di regressioni logistiche tra le frequenze ottenute dai soggetti per i 5 stili nelle prime 4 risposte e la scelta di un determinato stile nella quinta risposta (variabile dicotomica). Quest’analisi è stata ripetuta per tutti e 5 gli scenari.

L’analisi mostra risultati contrastanti. Come si evince dalla Tabella 9, solo in 2 dei 5 scenari di ENACT la regressione logistica ha fornito risultati significativi (con un $p \leq 0.05$) per tutti gli stili e questo significa che solo 2 dei 5 scenari presentano una coerenza tra le prime 4 domande e l’ultima. In altre parole solo nello scenario 1 e nello scenario 4 si può affermare che all’aumentare del valore di un determinato stile nei primi 4 item, aumenta la probabilità che questo stesso stile venga scelto all’item 5. Negli scenari 2, 3 e 5, questa coerenza non si verifica per tutti gli stili ma solo per alcuni. In particolare in questi tre scenari, la regressione su Avoiding e /o Compromising non è significativa.

SCENARIO 1	B	E.S.	Wald	df	Sig.	Exp(B)
Int	0,939	0,176	28,467	1	0,000	2,558
Dom	1,832	0,485	14,243	1	0,000	6,247
Obl	1,392	0,261	28,505	1	0,000	4,025
Avo	2,053	0,574	12,776	1	0,000	7,794
Com	0,945	0,349	7,348	1	0,007	2,572
SCENARIO 2	B	E.S.	Wald	df	Sig.	Exp(B)
Int	0,677	0,196	11,922	1	0,001	1,968
Dom	2,002	0,506	15,638	1	0,000	7,405
Obl	1,393	0,308	20,516	1	0,000	4,028
Avo	-0,450	0,992	0,206	1	0,650	0,638
Com	0,183	0,438	0,174	1	0,677	1,113
SCENARIO 3	B	E.S.	Wald	df	Sig.	Exp(B)
Int	0,914	0,175	27,294	1	0,000	2,495
Dom	1,988	0,411	23,411	1	0,000	7,3
Obl	1,319	0,3	19,369	1	0,000	3,741
Avo	0,878	0,339	6,715	1	0,010	2,406
Com	0,403	0,313	1,656	1	0,198	1,497
SCENARIO 4	B	E.S.	Wald	df	Sig.	Exp(B)
Int	0,64	0,229	7,793	1	0,005	1,896
Dom	2,319	0,519	19,959	1	0,000	10,163
Obl	1,543	0,268	33,008	1	0,000	4,677
Avo	1,187	0,36	10,893	1	0,001	3,278
Com	1,034	0,321	10,37	1	0,001	2,813
SCENARIO 5	B	E.S.	Wald	df	Sig.	Exp(B)
Int	0,974	0,245	15,779	1	0,000	2,648
Dom	1,712	0,345	24,575	1	0,000	5,541
Obl	1,524	0,27	31,912	1	0,000	4,591
Avo	0,665	0,515	1,667	1	0,197	1,945
Com	1,461	0,391	13,974	1	0,000	4,309

Tabella 9: Risultati della regressione logistica sui dati dei 5 scenari di ENACT

Conclusioni analisi 2: quest'analisi evidenzia alcune criticità sulla validità dell'item di controllo in 3 dei 5 scenari di ENACT. In particolare, si osserva che per gli stili **Avoiding e Compromising**, in questi scenari, **non si ottengono risultati significati**. Da questo possiamo solo concludere che in questi casi, all'aumentare della frequenza di questi stili nel primo blocco di item, non si evidenzia un aumento degli stessi sull'item di controllo. **Le ragioni di questa criticità molto probabilmente sono legate alla costruzione degli item stessi, che quindi andrebbero rivisti nella stesura originale.**

3.3.5 Analisi 3

“Studio della relazione tra stili e scenario con l’Analisi delle Corrispondenze”

D: Gli scenari sono caratterizzati da uno stile di risposta predominante o sono stati costruiti senza influenze stilistiche?

Finora abbiamo verificato che: 1) i primi 4 item all’interno di ciascuno scenario sono indipendenti tra loro, e che 2) l’item di controllo di ciascuno scenario, il n.5, non è sempre costruito in maniera coerente con quelli che lo precedono.

A questo punto ci si è chiesti se esiste una relazione tra stili e scenari. Sappiamo che durante la prova di assessment, alla fine di ogni scenario, il software indica lo stile maggiormente adatto in quella situazione e con quel tipo di bot, definiti sulla base della teoria di Rahim, secondo lo schema riportato di seguito:

n. scenario	Titolo	Stile bot	Stile “consigliato”
1	Sport	Obliging	Dominating
2	Motorbike	Integrating	Compromising
3	Restourant	Compromising	Avoiding
4	CD	Avoiding	Obliging
5	TV Programme	Dominating	Integrating

Tabella 10: “Stili consigliati” per ciascuno scenario di ENACT

Tenendo conto di questa scelta “consigliata”, alla fine di ogni sessione di gioco il software effettua un debriefing, mostrando le scelte effettuate dai soggetti e fornendo una valutazione rispetto a quanto sarebbe stato opportuno fare.

L’interesse si è focalizzato proprio sulla risposta n. 5, cercando di rispondere alle seguenti domande: i) i soggetti che hanno mostrato uno stile prevalente nelle prime 4 interazioni, lo hanno poi confermato nella quinta? ii) In caso affermativo, la risposta 5 era quella “suggerita” dalla teoria? iii) In caso contrario, quali sono le tendenze che emergono?

È importante sottolineare che l’item 5 non aveva originariamente uno status speciale rispetto alle altre, ed è stata interpretata come domanda di controllo solo in questa fase della ricerca, a seguito delle differenze emerse nelle analisi prima descritte. Una descrizione più dettagliata di quanto detto è riportata nel paragrafo 3.2.6.1.

In ogni caso, per rispondere al quesito n. 3, il primo passo è stato verificare le percentuali di soggetti che hanno mostrato uno stile prevalente nelle prime 4 risposte,

in base al criterio di 2 o più risposte sullo stesso stile. Come si evince dalla tabella che segue, in media il 61% dei soggetti ha un stile predominante. Di questo però solo in media il 32% circa ha effettuato la simulazione teoricamente corretta, anche se in alcuni scenari questa percentuale è particolarmente bassa (9.4% e 9.9%).

	Scen 1 SPORT (B=Obl) Stile Cons: DOM	Scen 2 MOTORBIKE (B=Int) Stile Cons: COM	Scen 3 RESTOURANT (B=Com) Stile Cons: AVO	Scen 4 CD (B=Avo) Stile Cons: Obl	Scen 5 TV PROGRAMME (B=DOM) Stile cons: INT	Media
% ss con 2 o + risp uguali all'ultima	65,9%	62,8%	62,8%	56,6%	57,4%	61,1%
di cui con stile "Consigliato"	9,4%	17,3%	9,9%	82,2%	40,5%	31,9%

Tabella 11: Percentuale di scelta di uno stile prevalente

Questo ci porta a chiederci se e quanto ciascuno scenario spinge i soggetti ad adottare un determinato stile di gestione del conflitto e, nel caso, se questo è corretto da un punto di vista teorico. L'introduzione di specifici parametri quali il tempo a disposizione e l'importanza del raggiungimento dell'obiettivo sia per il giocatore che per i bot, dovrebbe indicare al giocatore l'utilizzo di uno specifico stile di gestione. Tuttavia questo è sicuramente vero se si presuppone una fase preliminare di formazione ai soggetti sulla negoziazione e sul conflitto secondo le teorie di Rahim, come avviene nel caso di un percorso di formazione e valutazione della competenza con ENACT Training, mentre è meno probabile nel caso di assenza di informazioni sul tema, come nel nostro caso, in cui si è utilizzata solo la versione ASSESSMENT.

La tabella seguente mostra le tendenze espresse dal solo campione selezionato, ossia dai soggetti che hanno mostrato una tendenza specifica. In giallo sono indicate gli stili teoricamente corretti e in rosso le modalità principali scelte dai soggetti per ciascuno scenario. È chiaro che solo nello scenario 4 i due aspetti coincidono.

	Scen 1 SPORT (B=Obl) Stile Cons: DOM	Scen 2 MOTORBIKE (B=Int) Stile Cons: COM	Scen 3 RESTOURANT (B=Com) Stile Cons: AVO	Scen 4 CD (B=Avo) Stile Cons: Obl	Scen 5 TV PROGRAMME (B=DOM) Stile cons: INT
Avo	4,7%	0,0%	2,5%	6,8%	2,7%
Com	5,9%	1,2%	2,5%	8,2%	14,9%
Dom	4,7%	4,9%	9,9%	9,6%	16,2%
Int	57,6%	75,3%	74,1%	19,2%	16,2%
Obl	27,1%	18,5%	11,1%	56,2%	50,0%

Tabella 12: Incrocio dello stile consigliato con lo stile prescelto dal campione

Da questa prima analisi, si evince che gli scenari sembrano spingere i soggetti in una determinata direzione, che però nella maggior parte dei casi non è quella che ci si aspetta in base ai costrutti teorici di riferimento. Considerando la premessa fatta precedentemente, per cui i soggetti sottoposti al test, non erano formati sulle teorie di riferimento, non sorprende che l'assessment sia discordante dalla teoria. La correttezza di questo aspetto del gioco dovrebbe essere testata con persone che sono a conoscenza della teoria, o dopo una fase di training degli utenti.

In ogni caso, per averne conferma, si è fatto ricorso ad un'analisi statistica multivariata più sofisticata che rientra nell'ambito dell'analisi fattoriale, l'analisi delle corrispondenze (AC). Sul nostro campione, l'AC è stata utilizzata per analizzare il posizionamento dei diversi stili tra i diversi scenari. In altre parole, si vuole verificare se gli item proposti nei diversi scenari sono legati ad un particolare stile, oppure se sono indipendenti.

Come primo passo, è stata calcolata una matrice di dati che mostra il numero di risposte ai vari stili nei diversi scenari. La tabella, e ancora di più il grafico della Figura 17, mostrano i dati di partenza per l'AC. Come si può vedere, i vari stili presentano un andamento molto diverso. I due più scelti sono Integrating e Obliging, molto inferiori i valori degli altri tre. Ci siamo chiesti se gli scenari, così come sono stati ideati spingono i soggetti a rispondere con un certo stile.

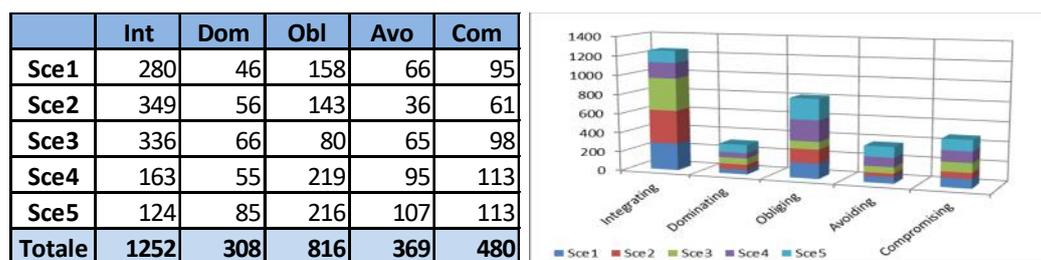


Figura 17: matrice e grafico dei dati per l'AC

Come si evince dalla Figura 18, dall'AC emerge un primo grande fattore che spiega gran parte della variabilità, circa l'88%, che cumulato con il secondo arrivano a coprire il 97 % di tutta la variabilità.

	F1	F2	F3	F4
Eigenvalue	0,088	0,009	0,003	0,000
Inertia (%)	87,942	9,052	3,006	0,000
Cumulative %	87,942	96,994	100,000	100,000

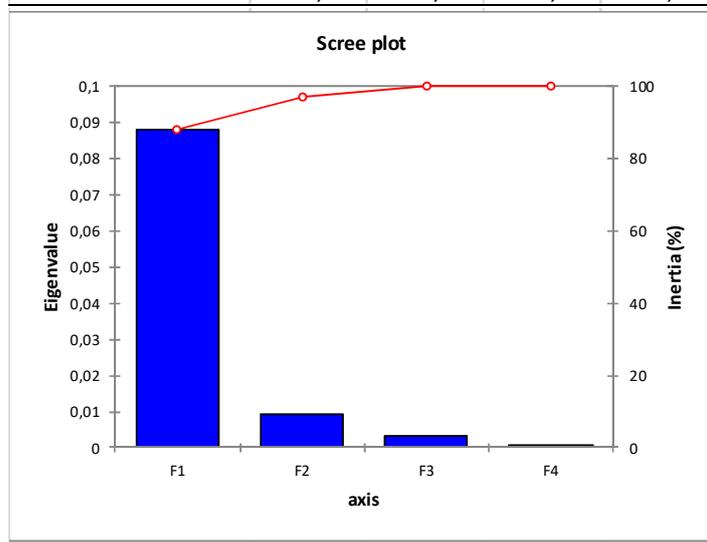


Figura 18: Dipendenza tra scenari e stili: AC

Nel piano costruito dal primo e secondo fattore, vedi Figura 19, si evincono le connessioni tra i diversi stili e i diversi scenari. Gli scenari che si posizionano in prossimità del baricentro (o incrocio degli assi) non sono caratterizzati da nessuno stile.

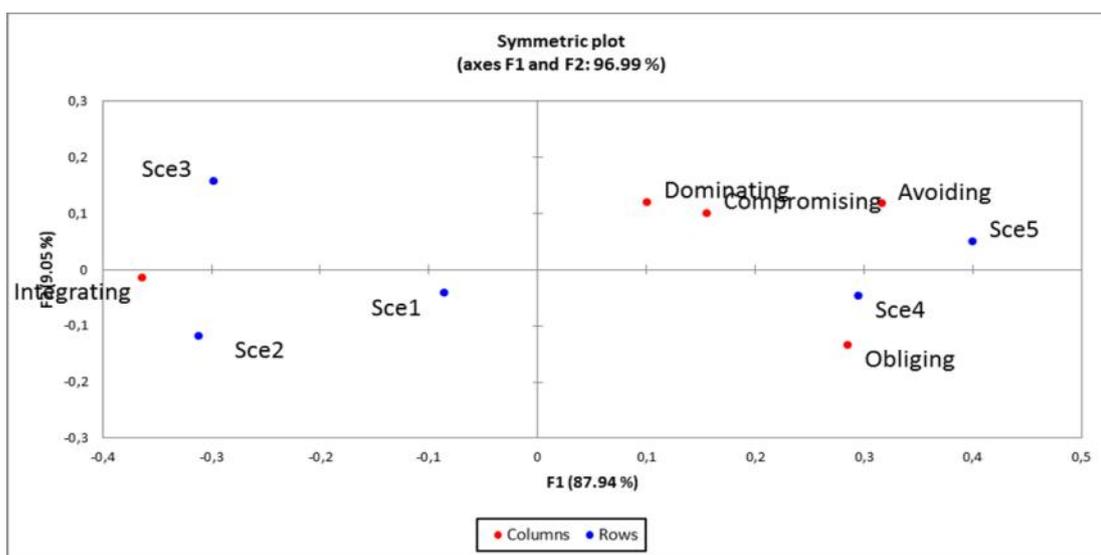


Figura 19: Rappresentazione dei risultati dell'AC

Nel nostro caso, è facile osservare che solo lo scenario 1 non presenta stili predominanti, mentre tutti gli altri sono influenzati almeno da uno stile di risposta. In particolare: gli scenari 2 e 3, sono influenzati dallo stile Integreting, lo scenario 4 è vicinissimo ad Obliging mentre il quinto, è caratterizzato da Avoiding. È necessario ricordare che i dati considerati in questa analisi coinvolgono l'intero campione, e non il sottoinsieme di soggetti con un stile di risposta predominante, come nell'analisi precedente.

Se però si confrontano questi risultati con quelli del paragrafo precedente, si evidenzia un dato interessante: in tre casi, l'AC conferma quanto detto prima, infatti gli scenari 2 e 3 risultano influenzati dallo stile Integreting e lo scenario 4 da Obliging. In maniera più dettagliata possiamo osservare che:

- 1) Nell'AC, lo scenario 1 non risulta legato a nessuno stile specifico, essendo abbastanza vicino al centro degli assi, ma il campione mostra una tendenza ad adottare uno stile Integreting (57,6%), invece di Dominating (4,7%) stile consigliato, che tra l'altro è quello meno scelto dal sottoinsieme dei soggetti. Questo dato ci fa supporre che questo scenario è teoricamente costruito in senza influenze dirette e non fa emergere nessuno stile specifico, tantomeno quello considerato più adatto alla situazione.

Nota: nelle interviste successive al testing, molti soggetti riportavano di aver avuto difficoltà a mantenere uno stile coerente in quanto la simulazione appariva abbastanza irrealistica, ed è per questo che spesso tendevano a cercare una soluzione integrata.

- 2) Nell'AC, gli scenari 2 e 3 sono in molto vicini allo stile Integreting e questo è confermato anche dal campione. Infatti in entrambi i casi, il 62,8% dei soggetti del sottocampione, ha mantenuto uno stile specifico e di questi rispettivamente il 75,3% e il 74,1% hanno scelto proprio Integreting. Lo stile consigliato, Compromising, è stato scelto solo dallo 1,2% dei soggetti. Questo ci dice che gli scenari tendono verso uno stile specifico e per di più non è quello "consigliato".

Nota: i ricercatori che si sono occupati della costruzione degli item hanno riportato una notevole difficoltà a creare delle differenze tra gli stili Integreting e Compromising, dato che nella linguaggio parlato i due stili si sovrappongono molto e

che la loro differenza risiede più nel semantico che non nell'uso corrente (vedi paragrafo 3.2.6.1).

- 3) Lo scenario 4 è quello maggiormente chiaro. Lo stile consigliato e il trend adottato dai soggetti, nella maggior parte dei casi, coincidono. Il 56,6% del campione ha utilizzato uno stile principale e di questi il 82,2% ha adottato proprio lo stile Obliging. L'AC conferma questo dato, mostrando una notevole vicinanza tra queste due variabili. Questo ci suggerisce che lo scenario ha Obliging come stile predominante.
- 4) Lo scenario 5 invece è quello più anomalo: l'AC infatti lo mostra molto influenzato dallo stile Avoiding, sebbene una percentuale molto bassa di soggetti del sotto campione ha espresso questa tendenza comportamentale (solo il 2,7% dei soggetti). Questo ci suggerisce che i punteggi dell'intero campione sono molto variegati e che lo stile Avoiding è stato scelto spesso dai soggetti che non hanno manifestato uno stile specifico.

Conclusioni analisi 3: incrociando i risultati ottenuti analizzando gli stili di gestione scelti dai soggetti nei diversi scenari, con i risultati dell'AC, possiamo affermare che **gli scenari sono “caratterizzati” da uno stile specifico, che quasi sempre non coincide con quello “suggerito” in base al costrutto teorico.** In altre parole, solo uno scenario, il primo, sembra essere libero dalla predominanza di uno stile, mentre negli altri 4 scenari, non solo si evince uno stile predominante, ma spesso (3 volte su 4) questo stile non è neanche quello teoricamente “corretto” secondo i costrutti teorici con cui si è costruito lo strumento. Se a questi dati si aggiungono le riflessioni dei soggetti e dei ricercatori che hanno partecipato allo studio, appare evidente la necessità di intervenire sugli item che compongono i singoli scenari e modificarli per renderli imparziali, cioè non identificato da uno specifico stile.

3.3.6 Analisi 4

“Studio della relazione tra ROCI e ENACT mediante l’Analisi delle Componenti Principali (ACP) e l’Analisi delle Correlazioni Canoniche (ACC)”

D: Gli stili misurati dai due strumenti e più in generali i due strumenti stessi hanno una struttura latente simile? Misurano gli stessi concetti? Come possono essere confrontati?

Ultimo passo di quest’analisi esplorativa è caratterizzato dal tentativo di identificare un metodo statistico che ci permettesse di mettere a confronto i due test, che come ricordiamo sono molto differenti tra loro. Come si evince dalla teoria di Rahim, il ROCI II ha una struttura fattoriale a 5 fattori, uno per ogni stile, definito dai punteggi ottenuti su un gruppo di item indipendenti per ciascuno stile, misurati su scala Likert. ENACT, al contrario, presenta un dataset molto complesso e articolato, con frequenze raggruppabili sia per stile che per scenario, quest’ultimo oltretutto del tutto assente nel ROCI II.

Il primo passo è stato di utilizzare l’analisi delle componenti principali (ACP) sui dati di ENACT al fine di estrarre una componente principale (fattore) per ogni stile per poi effettuare un confronto con quelli del ROCI.

Nel nostro caso, il primo passo è stato calcolare le componenti principali per ciascuno stile di ENACT e successivamente confrontarle con quelle calcolate sul ROCI. Purtroppo quest’analisi non ha dato risultati interessanti, poiché in nessun caso si è ottenuto un fattore unico che spiegasse la maggior parte della variabilità di uno dato stile. Più nello specifico (vedi Figura 20), l’analisi mostra la necessità di estrarre almeno due fattori, per spiegare poco più del 50% della variabilità dei dati ed inoltre si evidenzia un andamento simile per ciascuno stile e cioè il valore del primo fattore è sempre intorno al 1.5, mentre il 2 e il 3 si aggirano intorno all’1.0. Questo rende ancor più difficile la scelta del numero di fattori da considerare.

Nonostante la non conclusività di questo risultato, si è scelto di mantenere il primo fattore di ciascuno stile di ENACT, in quanto comunque conservano la maggior parte dell’informazione. Successivamente è stata utilizzata un’analisi statistica multivariata per analizzare la struttura di correlazione tra i fattori dei due test: l’Analisi delle Correlazioni Canoniche (ACC).

Stile Compromising	F1	F2	F3	F4	F5
Eigenvalue	1,58	1,05	1,00	0,88	0,48
Variability (%)	31,65	21,04	20,02	17,68	9,61
Cumulative %	31,65	52,69	72,71	90,39	100,00
Stile Avoiding	F1	F2	F3	F4	F5
Eigenvalue	1,74	1,06	0,91	0,85	0,44
Variability (%)	34,86	21,21	18,19	17,04	8,70
Cumulative %	34,86	56,07	74,26	91,30	100,00
Stile Obliging	F1	F2	F3	F4	F5
Eigenvalue	1,76	1,10	0,86	0,67	0,62
Variability (%)	35,25	21,92	17,18	13,34	12,31
Cumulative %	35,25	57,17	74,34	87,69	100,00
Stile Dominating	F1	F2	F3	F4	F5
Eigenvalue	1,89	1,02	0,89	0,72	0,49
Variability (%)	37,73	20,49	17,71	14,35	9,71
Cumulative %	37,73	58,23	75,93	90,29	100,00
Stile Integrating	F1	F2	F3	F4	F5
Eigenvalue	1,76	1,01	0,96	0,72	0,56
Variability (%)	35,11	20,13	19,19	14,44	11,12
Cumulative %	35,11	55,25	74,44	88,88	100,00

Figura 20: ACP sui dati ENACT

Come mostra la Figura 21, la prima correlazione canonica, ossia la correlazione tra la prima componente canonica del blocco ENACT e la prima componente canonica del blocco ROCI, è uguale a 0.52.

> res\$cor		[1] 0.51609910	
> as.matrix(ROCI)		as.matrix(ENACT)	
	Xcan1		Ycan1
INT	1.14653653	Integrating	0.46725721
DOM	-0.16866016	Dominating	-0.40765829
OBL	0.24997101	Obliging	0.41923287
AVO	0.03925166	Avoiding	0.33603424
COMP	0.17811402	Compromising	0.01181954

Figura 21: Valori della prima coppia di componenti canoniche sui due test

A questo punto è interessante analizzare quanto e come le variabili dei due blocchi hanno contribuito a costruire questa coppia di componenti. In Figura 21 sono riportati i pesi di ciascuna variabile sulla prima componente canonica. Si può notare che le componenti dei due test hanno una struttura simile, in termini di pesi. In particolare, per entrambi i fattori le variabili con un peso maggiore in ordine decrescente sono Integrating, Obliging e Dominating. Inoltre, in entrambi i casi quest'ultima ha un peso negativo.

Questo risultato è di grande interesse e per nulla scontato in quanto ci indica che le due strutture latenti in realtà sono molto simili, caratterizzate cioè dagli stessi stili e dalle stesse direzioni. Infatti, in entrambi i casi Dominating, va in senso inverso rispetto a tutte le altre variabili.

Il limite di quest'analisi non risiede nel metodo bensì nella rappresentatività dei dati. È chiaro infatti che non ottenendo buoni risultati con l'ACP, le componenti principali spiegano poca variabilità, e che quindi la correlazione della prima correlazione canonica non può risultare altissima.

Conclusioni analisi 4: Con quest'analisi abbiamo cercato di trovare una metodologia psicométrica che ci permettesse di confrontare due strumenti, che seppur nati dagli stessi assunti teorici, di fatto misurano e valutano gli stessi costrutti in maniera completamente diversa, diversità che si esprime appieno nel tipo di dati raccolti nelle sperimentazioni.

I risultati mostrano che l'Analisi delle Componenti principali e l'analisi delle correlazioni canoniche possono rappresentare un buon metodo, in quanto ci permettono dapprima di quantificare statisticamente i dati di ENACT e successivamente di mettere a confronto le strutture latenti dei due strumenti, misurando non le differenze ma le parti comuni.

Nel nostro caso, **l'analisi delle corrispondenze canoniche mostra che** questi due strumenti seppur nella forma sono completamente diversi, **presentano una struttura latente molto simile**, in altre parole, misurano le stesse cose, con le stesse modalità, con particolare riferimento agli stili Integreting, Dominating e Obliging. È interessante notare che questi tre stili sono gli stessi che mostravano buone correlazioni anche nelle analisi di fine progetto.

Infine, anche con l'ACC emerge **la poca significatività di Avoiding e Compromising**, così come nelle analisi precedenti e in quelle di fine progetto, e questo **ci fa ritenere che gli item utilizzati in ENACT per questi due stili non sono rappresentativi**. Si suppone, che una volta modificati gli item, l'ACP e di conseguenza l'ACC forniranno dati più robusti e maggiormente significativi

In conclusione, nonostante i limiti oggettivi dei dati, possiamo affermare che il binomio, ACP e ACC, può rappresentare un buon metodo di confronto tra questi due test, completamente diversi tra loro, fornendo le basi per effettuare degli studi più ampi e completi.

4 Conclusioni

Questa ricerca, risultato finale del dottorato in Human Mind and Gender Studies, si presenta come un lavoro multidisciplinare inteso ad integrare tre temi fortemente attuali: assessment, soft skill e serious game, all'interno di un unico campo d'indagine: la valutazione delle competenze individuali attraverso tecniche e strumenti tecnologicamente avanzati.

Operativamente questo si è tradotto in un duplice obiettivo: da una parte *proporre ENACT, serious game di ultima generazione, come strumento di testing alternativo ai tradizionali metodi "Carta e Matita"*; dall'altra *presentare i risultati dell'analisi esplorativa condotta sui dati ottenuti durante il progetto e finalizzata all'identificazione di un metodo psicometrico statisticamente valido per il confronto dei due test, e l'individuazione degli elementi critici necessari di ulteriori studi e approfondimenti.*

Creare un serious game evoluto, in grado cioè di effettuare una valutazione delle competenze, durante il gioco, che risulti accattivante, divertente e allo stesso tempo valido e attendibile, ma soprattutto equivalente al test originale da cui ha preso origine, è un processo lungo, complesso, articolato e costoso.

Come abbiamo visto, nel corso dei primi due anni, il consorzio europeo, formato da enti, fondazioni e piccole e medie imprese, è riuscito a realizzare uno strumento alternativo dalle caratteristiche del tutto peculiari e innovative, partendo da un test di 28 domande quale il ROCI II. Le difficoltà proprie della progettazione e dello sviluppo di un reattivo sotto forma di videogame, che mantenesse i criteri del test di partenza ma allo stesso tempo abbracciasse tutti i criteri del gaming, hanno richiesto uno sforzo notevole, soprattutto perché realizzato in un ambito di ricerca, con esigui fondi a disposizione e lontano quindi dalle logiche industriali e commerciali, che hanno fatto dei video game un settore in continua e rapida espansione.

L'analisi esplorativa condotta nell'ultimo anno, dimostra infatti che, se da una parte, il test possiede tutte le caratteristiche per essere una misura valida e attendibile della capacità di gestire i conflitti interpersonali, così come definiti e misurati da Rahim, dall'altra evidenza che, probabilmente, lo sforzo realizzato nella prima fase di

progetto sia andato a discapito della fase successive determinando pertanto alcune criticità e problematiche.

Più nello specifico, l'analisi ha fornito due tipi di informazione, la prima prettamente legata alla validità e all'attendibilità di ENACT, che possono essere potenziate, in maniera relativamente semplice, attraverso una serie di aggiustamenti alla metodologia sia di sviluppo che di testing, la seconda, più squisitamente psicometrica, è invece connessa alla possibilità di confrontare, secondo le regole della statistica classica, i due strumenti che sono nella forma e nella sostanza due prodotti completamente diversi.

Innanzitutto, per quanto riguarda la robustezza del nuovo strumento, tutte le analisi condotte portano a concludere che è necessario:

- 1) aumentare il numero dei partecipanti al test, al fine di ottenere un campione maggiormente rappresentativo e quindi migliorarne l'attendibilità.
- 2) rivedere gli item di alcune scenari, in particolare Avoiding e Compromising, sia sul gruppo di item principale, sia sull'item di controllo, al fine di rendere il test più realistico e quindi migliorarne la validità.

Inoltre per i successivi lavori, è fondamentale capire come gestire il dato relativo allo stile consigliato. Come abbiamo visto infatti, questo elemento ha un diverso impatto a seconda che la valutazione venga effettuata alla fine di un percorso di formazione, dove tutti i soggetti hanno una visione chiara delle teorie di riferimento e quindi una maggiore consapevolezza delle proprie azioni nei diversi contesti di conflitto "virtuali" in cui sono coinvolti, oppure come momento di valutazione a se stante, fotografia cioè delle competenze specifiche sviluppate dai soggetti fino ad allora. Appare evidente, in questo secondo caso, che il concetto di scelta consigliata decade, se non come elemento di confronto per avviare un successivo percorso formativo teso a migliorare queste stesse capacità.

Infine, l'ADE è stata utilizzata per verificare se effettivamente i due strumenti misuravano le stesse cose. Non potendo effettuare questa validazione in maniera diretta, date le notevoli differenze, si è reso necessario indagare le strutture latenti,

con due tecniche proprie della teoria classica dei test: l'analisi delle componenti principali e l'analisi delle correlazioni canoniche.

L'uso combinato di queste due tecniche si è dimostrato l'unico in grado di mettere a confronto i due strumenti e offre ottime speranze sul futuro di questi lavori, in quanto nonostante i limiti evidenziati e legati soprattutto alle criticità sopra esposte, mostrano l'esistenza di un buon parallelismo tra i due strumenti, che sembrano misurare le stesse cose e nello stesso modo, soprattutto per gli stili Integrating, Dominating e Obliging, gli stessi che hanno mostrato valori significativi nelle analisi di fine progetto.

Questo risultato, che potrebbe sembrare scontato, in realtà ha un grosso valore in termini psicometrici, in quanto ci conferma che la strada intrapresa è quella giusta e che migliorando i pochi elementi di criticità metodologica sopra descritti, è possibile effettuare un confronto valido e attendibile tra due test che in comune hanno solo l'oggetto di valutazione. Sebbene ad oggi, questo potrebbe apparire fine a se stesso, non lo è in una prospettiva a medio e lungo raggio, caratterizzata da un costante e crescente interesse verso strumenti di valutazione sempre più evoluti, che sfruttano al massimo le innovazioni tecnologiche e cercano di mantenere il passo con i cambiamenti di questo secolo, con il risultato di un allontanamento da quelle che possono ormai essere considerate le tecniche tradizionali di misurazione delle soft skill.

Ed è per questo motivo, in conclusione, che il test qui presentato, ENACT, può essere considerato uno strumento pionieristico, sicuramente all'avanguardia, nonostante le criticità e le debolezze che in questa sua prima versione mostra.

Allegato 1 - ROCI II Form C

Cerca di ricordare le più recenti situazioni di conflitto in cui sei stato coinvolto con dei tuoi pari e indica quanto sei d'accordo con le affermazioni che seguono, utilizzando la scala da 1 a 5 (per niente d'accordo – completamente d'accordo).

*Per niente
d'accordo*

*Completamente
d'accordo*

	1	2	3	4	5
1. Cerco di indagare su un problema con i miei colleghi per trovare una soluzione accettabile per tutti.					
2. Generalmente cerco di soddisfare le esigenze dei miei coetanei					
3. Cerco di evitare di essere "messo in difficoltà" e tendo a tenere il mio conflitto per me.					
4. Cerco di integrare le mie idee con quelle dei miei colleghi così da prendere una decisione comune.					
5. Cerco di lavorare con i miei colleghi per trovare una soluzione a un problema che soddisfa le aspettative di tutti.					
6. Con i miei coetanei, solitamente evito di discutere apertamente delle mie differenze					
7. Cerco di trovare un compromesso per risolvere le situazioni di stallo					
8. Uso la mia influenza per far sì che le mie idee siano accettate.					
9. Uso la mia autorità per prendere decisioni a mio favore.					
10. Di solito sono accomodante verso i desideri dei miei coetanei.					
11. Cedo ai desideri dei miei coetanei.					
12. Con i miei coetanei scambio accurate informazioni per risolvere insieme un problema.					
13. Di solito permetto concessioni ai miei coetanei.					
14. Di solito propongo una via di mezzo per risolvere una situazione di stallo.					
15. Negoziò con i miei coetanei in modo che un si possa trovare un compromesso					
16. Cerco di evitare disaccordi con miei coetanei					
17. Evito di litigare con i miei coetanei.					
18. Io uso la mia esperienza per prendere decisioni a mio favore					
19. Spesso seguo i suggerimenti dei miei coetanei.					
20. Io uso il "dare e avere", così da poter trovare un compromesso					
21. Generalmente in una discussione perseguo le mie ragioni					
22. Cerco di evidenziare tutte le preoccupazioni in modo che i problemi possano essere risolti nel miglior modo possibile.					
23. Collaboro con i miei colleghi per arrivare a decisioni accettabili per tutti.					
24. Cerco di soddisfare le aspettative dei miei coetanei.					
25. A volte uso il mio potere per vincere le competizioni.					
26. Cerco di nascondere il mio disaccordo ai miei coetanei così da evitare possibili rancori.					
27. Cerco di evitare spiacevoli scambi con i miei coetanei.					
28. Cerco di lavorare con i miei colleghi per una corretta comprensione dei problemi.					

Allegato 2 - ENACT ASSESSMENT – Screenshot Scenari

Introduzione al gioco



Scenario 1- Sport





Scenario 2- Motorbike





Scenario 3 – Restaurant

enactgame (more info on enactgame.eu >>)

Pizza o ristorante cinese?



Roberta
Tu sei Roberta, e desideri ardentemente mangiare la pizza. E' un desiderio troppo forte!



Vittoria
Vittoria è appassionata di cucina cinese e adora mangiare cibi orientali con le bacchette.

Back Next

enactgame (more info on enactgame.eu >>)

Pizza o ristorante cinese?

Dove andrete a mangiare questa sera?



Roberta



Vittoria

Ricorda che:
 - Non sai ancora dove andare dopo cena, e vorresti non doverne avere troppo.
 - Non vorresti litigare con la tua amica soltanto per il cibo.

Back Next

Time



Importance

for you ★★★★★ for other ★★★★★

enactgame (more info on enactgame.eu >>)

Ti trovi nel mezzo della discussione con Vittoria.

Start



enactgame (more info on enactgame.eu >>)

Che fare? Non vedo l'ora di uscire! Vieni qualcun altro oltre a noi?

Non ne sono ancora sicura, ho anche un po' fame (incredibile però!)

Point to previous Press to select



enactgame (more info on enactgame.eu >>)

Il turno è: pizza o cinese?

Io avrei voglia di pizza, ma so che tu preferisci mangiare orientale... vero?

Mi hanno detto che c'è una nuova pizzeria che fa la pizza buonissima... Potremmo mangiare orientale la prossima volta?

Mi piacerebbe andare alla nuova pizzeria, mi hanno detto che non fanno solo pizza, ti va di vedere cosa preparano?

Abbiamo appena pranzato, come si può decidere cosa vogliamo stasera?

Mi hanno detto che c'è una nuova pizzeria che fa le pizze buonissime... Andiamo a provarla!



enactgame (more info on enactgame.eu >>)

Avrei voglia particolare di cinese stasera, possiamo provare la nuova pizzeria in settimana se ti va, che ne pensi?

Potremmo andare a mangiare al ristorante messicano questa sera. Mi sembra un buon compromesso!

Se è la tua così tanto il cinese, per questa volta possiamo andare lì!

Proviamo a vedere su internet se c'è un posto dove fanno sia pizza che cibo orientale?

Io trovo quelle bacchette impossibili da usare! Meglio la pizza!

Ma non sappiamo ancora cosa fare dopo... Dovremmo aspettare per decidere!



enactgame (more info on enactgame.eu >>)

Ne conosce uno, poco fuori dal centro. C'è ne dici?

Non abbiamo la macchina stasera, e la pizzeria è la più vicina da qui.

E la macchina? Aspettiamo di sapere cosa faremo dopo, o no ripensiamo?

Non abbiamo la macchina purtroppo. Ci andiamo la prossima settimana, ed oggi andiamo in pizzeria?

So che ti tieni a mi farebbe piacere, ma non abbiamo la macchina per arrivarci...

Sarebbe ottimo, ma non abbiamo la macchina per arrivarci, come possiamo fare?



enactgame (more info on enactgame.eu >>)

Ideali! Ci sarebbe un takeaway che ci porta sia pizza che cibo cinese. E poi potremmo uscire. Che ne pensi?

Se questa scelta fa felice te, allora mi sta bene! Questa sera vorrei che tu fossi soddisfatta in ogni caso!

Potrebbero metterci troppo tempo per la consegna... Guardiamo un po' di TV e aspettiamo che ci venga fatto.

E se aspettiamo troppo tempo? Dice che un altro tipo di cucina stasera sarebbe il miglior compromesso!

Ottima opzione! Prima di decidere, potremmo anche dare un'altra occhiata ai tipi di cucine ai ristoranti nei dintorni!

Non mi va di rimanere a casa... Con tutte le volte che siamo andati al cinese stasera davvero accontentarmi!





Scenario 4 – CD





Scenario 5- TV Programme





Esempi di Schede di Debriefing



Bibliografia

- Abt C. (1970), *Serious Games*, New York, Viking Press.
- Anderson, L., Krathwohl, R., Alrasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., & Wittrock, M. (Eds.) (2001). *Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy* New York, NY: Longman.
- Angehrn A., Maxwell K., Luccini M., Rajola F., (2009), *Designing Effective Collaboration, Learning and Innovation Systems for Education Professionals*, *International Journal of Knowledge and Learning (IJKL)*, Vol. 5 No.2.
- Arnab S., Um T., Carvalho M. B., Bellotti F., de Freitas S., Louchart S., Suttie N., Berta R. and De Gloria A., "Mapping Learning and Game Mechanics for Serious Games Analysis", *British Journal of Educational Technology*.
- Baalsrud Hauge J., Belletti F., Berta R., Carvalho M. B., De Gloria A., Lavagnino E., Nadolski R., Ott M., (2013) *Field assessment of Serious Games for Entrepreneurship in Higher Education*, *Journal of Convergence Information Technology*, pp. 1-12, Number 3.
- Bachen, C. M., Hernandez-Ramos, P. F., & Raphael, C. (2012). *Simulating REAL LIVES: Promoting global empathy and interest in learning through simulation games*. *Simulation & Gaming*, 43(4) · See more at: <http://www.scu.edu/profiles/?p=3433#sthash.WFuJ4kba.dpuf>
- Bastos N. P., Moran Sanchez-Cantón B., Hervello C. (2012), *ENTRExplorer: a Serious Game for Immersive Entrepreneurs*, *Atas do Encontro sobre Jogos e Mobile Learning*. Braga: CIED.
- Belletti F., Berta R., De Gloria A., Primavera L, (2009) "Adaptive Experience Engine for Serious Games", *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 1, No.4.
- Belletti, F., Kapralos, B., Lee, K., Moreno-Ger, P., Berta, R., (2013), "Assessment in and of Serious Games: An Overview", *Hindawi Advances in Human-Computer Interaction*.
- Bellotti F., Berta R., De Gloria A., (2010), "Designing Effective Serious Games: Opportunities and Challenges for Research", *Special Issue: Creative Learning with Serious Games*, *Int'l Journal of Emerging Technologies in Learning (IJET)*, Special Issue: Creative Learning with Serious Games Vol.5, pp. 22-35.
- Bellotti F., Berta R., De Gloria A., D'Ursi A., and Fiore V. (2012). *A serious game mode! far cultura! heritage*. *ACM J. Comput. Cult. Herit.* 5, 4.
- Bellotti F., Berta R., De Gloria A., Lavagnino E., Antonaci A., Dagnino F., Ott M., (2013), "Gamified Short Course for Promoting Entrepreneurship among ICT Engineering Students", *Proceedings of IEEE International Conference on Advanced Learning Technologies (ICALT)*, Beijing, China.
- Bellotti F., Berta R., Ferretti E., DeGloria A., Margarone M., (2003) "VeGame: Field Exploration of Art and History in Venice", *IEEE Computer*, special issue on *Handheld Computing*, Vol. 36, No. 9. pp. 48-55.

- Bente G., J. Breuer (2009) Making the implicit explicit: Embedded measurement in serious games. In *Serious Games: Mechanisms and Effects*. U. Ritterfeld, M. J. Cody, P. Vorderer (Eds.), Mahwah, NJ: Routledge, Taylor and Francis.
- Berta R., Bellotti F., De Gloria A., Pranantha D., Schatten C., (2013) "Electroencephalogram and Physiological Signal Analysis for Assessing Flow in Games", *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 5, No.2, 567-578, 164-175.
- Bisiacchi, P. S.; Mapelli, D.; Mondini, S.; Vestri, A. (2003) *Esame neuropsicologico breve*, Milano, Raffaello Cortina Editore.
- Bolasco S, (6^a ristampa 2014) *Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*, 1999, Roma, Carocci.
- Boncori, L. (1993) *Teoria e tecniche dei test*, Torino, Bollati Boringhieri
- Boncori, L. (2006) *I Test in Psicologia*, Bologna, Il Mulino
- Chen J., (2007) "Flow in games (and everything else)" *Communications of the ACM* 50, 4, pp. 31-34.
- Chen M., Johnson S., (2004) "Measuring flow in a computer game simulating a foreign language environment," Available online at: http://www.marckdangerchen.net/pub/flow_in_game_simulating_flow.pdf
- Chiang Y. T., Cheng C. Y., Un S. S. J., (2008) "The Effects of Digital Games on Undergraduate Players' Flow Experiences and Affect," in proceedings of the Second IEEE International Conference on Digital Games and Intelligent Toy Enhanced Learning (DIGITEL '08)
- Chi-Hsiung Liang, (2012) "Solving family communication problems between children and parents by using Mobile Serious Games," *Interactive Collaborative Learning (ICL)*, 2012 15th International Conference on, vol., no., pp.1,6.
- Cole, S.W., Yoo, D.J., Knutson, B. (2012). Interactivity and Reward-Related Neural Activation During a Serious Videogame. *PLoS ONE* 7(3).
- Connolly, T. M., E. A. Boyle, E. MacArthur, T. Hainey, J. M. Boyle (2012). "A systematic literature review of the empirical evidence on computer games and serious games." *Computers and Education* 59(2), pp. 661-686.
- Connolly, T.M., Boyle, E.A., Stansfield, M.H., Hainey, T., (2007) "The Potential of Online Games as a Collaborative Learning Environment", *Journal of Advanced Technology for Learning*.
- Csikszentmihalyi M., (1990) *Flow: The Psychology of Optimal Experience*, New York: Harper & Row.
- Del Bianco A., Serrano-Laguna A., Freire M., Martfnez-Ortiz., Fernandez-Manjon B., (2013) E-learning Standards and learning Analytics: Can Data Collection Be Improved by Using Standard Data Models?. *IEEE Engineering Education Conference (EDUCON)*, pp 1255-1261, Berlin, Germany.
- Del Bianco, A., Torrente, J., Moreno-Ger, P., Fernandez-Manjon, B., (2010) Integrating adaptive games in student-centered virtual learning environments.

International Journal of Distance Education Technologies, Volume 8, Issue 3, Pages 1-15.

Dell'Aquila, E., Pacella, D., Marocco, D. (2014) "Il Progetto ENACT: nuove tecnologie per la formazione alla negoziazione". In Airenti G., Cruciani M., Di Nuovo S., Perconti P., Plebe A., In: Le scienze cognitive a confronto. Oltre i confini della teoria, pp 162-181. Corisco Edizioni.

Dell'Aquila, E., Pacella, D., Marocco, D. (2014) "Il Progetto ENACT: nuove tecnologie per la formazione alla negoziazione". in Airenti G., Cruciani M., Di Nuovo S., Perconti P., Plebe A., Le scienze cognitive a confronto. Oltre i confini della teoria, pp 162-181 Corisco Edizioni.

Dell'Aquila, E., Yasar Akyar, O., Marocco, D., (2015) ENACT: D4.1 Negotiation theories and psychological modelling, a literature review.

Dell'Aquila E & Marocco D. (2015). Vivere con gli altri: giochi di ruolo e competenze relazionali. In: Vita naturale, vita artificiale. Tecniche di simulazione e applicazioni educative e cliniche. A cura di S. Di Nuovo e A. Cangelosi, Franco Angeli.

Dell'Aquila E., Marocco D. (2015). Vivere con gli altri: giochi di ruolo e competenze relazionali. In: Vita naturale, vita artificiale. Tecniche di simulazione e applicazioni educative e cliniche. A cura di S. Di Nuovo e A. Cangelosi, Franco Angeli Editore.

Dell'Aquila E., Marocco D., Ponticorvo M., Di Ferdinando A., Schembri M., Miglino O. (2017) Educational Games for Soft-Skills Training in Digital Environments - New Perspectives. Springer-Verlag.

Dell'Aquila E., Yasar Akyar, O., Marocco D. (2014) ENACT: D4.3. Online contents on negotiation training and assessment.

Dell'Aquila, E., Marocco, D., Ponticorvo, M., di Ferdinando, A., Schembri, M., Miglino, O. (2016) "Educational Games for Soft-Skill Training in Digital Environments", Advances in Game-Based Learning, Springer International Publishing.

Djaouti D., Alvarez J., Jessel J.P., Methel G., Molinier P., (2008) "A Gameplay Definition through Videogame Classification", International Journal of Computer Game Technology, Hindawi Publishing Corporation, Quarter 1.

Djaouti D., Alvarez J., Jessel JP, (2011) "Classifying Serious Games: The G/P/S Model" IGI Global.

Doucet L., Srinivasan, (2010). Designing entertaining educational games using procedural rhetoric: a case study. In Proc.s 5th ACM SIGGRAPH Symp. on Video Games, Los Angeles, Ca.

Egenfeldt-Nielsen S., (2006) Overview of research on the educational use of videogames, Digitali Kompetanse, 3(1).

Ercolani A. Paola (a cura di) (2007). Strumenti statistici per la ricerca, la valutazione e la diagnosi in psicologia. Cortina Raffaello Editore.

Fabbris L., (1997). Statistica Multivariata. McGraw-Hill, Milano.

- Felicial P., (2011) "Handbook of Research on Improving Learning and Motivation through Educational Games: Multidisciplinary Approaches" , IGI Global.
- Ferdinando, A. D., Schembri, M., Ponticorvo, M. & Miglino, O (2015) Agent Based Modelling to Build Serious Games: The Learn to Lead Game. In Vicente, F., Manuel, J., Álvarez-Sánchez, Ramón, J., Paz López, Félix et al (editors), Bioinspired Computation in Artificial Systems. Springer.
- Gardner H, (2000) *Formae mentis*, Feltrinelli, Milano.
- Gee J. P., (2003) *What Video Games Have to Teach Us about Learning and Litcrocy*, New York: Palgrave Macmillan.
- Gigliotta, O., Marocco, D., Yasar Akyar, O. (2015) ENACT: D7.1 Report on pilot studies.
- Gigliotta, O., Marocco, D., Yasar Akyar, O. (2016) ENACT: D7.2 Report on large scale trials.
- Grandi L. G. (a cura di) (2010). *La psicodiagnosi*. Cortina (Torino) Editore.
- Greitzer F.L., Kuchar O.A., Huston K., (2007) "Cognitive Science Implications for Enhancing Training Effectiveness in a Serious Gaming Context", *ACM J. Educational Resources in Computing*, vol. 7, no. 3.
- Harrow A. J., (1972) "A taxonomy of the psychomotor domain: A guide for developing behavioral objectives", New York: McKay.
- Howell K., E. Glinert, L Holding, C. Swain, (2007) "How to build serious games." *COMMUNICATIONS OF THE ACM*, v. 50 issue 7, p. 44-49.
- Jarvis, S., de Freitas, S., (2009) Evaluation of a Serious Game to support Triage Training: In-game Feedback and its effect on Learning Transfer. *Proceedings of 2009 IEEE Conference in Games and Virtual Worlds for Serious Applications*.
- Kincaid, J.P., Westerlund, K.K., (2009) "Simulation in education and training," *Proceedings of the 2009 Winter Simulation Conference (WSC)*, Dec.
- Kirkpatrick, D. L., (1998) "Evaluating Training Programs: The Four Levels", Berrett-ehler Publishers, Inc, San Francisco.
- Kline, P. (1996). *Manuale di psicometria. Come costruire, valutare e applicare un test psicologico*. Astrolabio Ubaldini Editore.
- Kline, P., (2013) *Handbook of Psychological Testing (second edition)* Routledge.
- Klopfer E, Osterweil S., Salen K., (2009) *Moving learning Games Forward, Obstacles Opportunities & Openness*, The Education Arcade, Massachusetts Institute of Technology, Available online at: http://education.mit.edu/papers/MovinglearningGamesForward_EdArcade.pdf.
- Krathwohl R., Bloom B.S., Masia B.B, (1964) *Taxonomy of educational objectives: The classification of educational goals. Handbook II: The affective domain*. David McKay Company.
- Lee, C. (1983). Self-efficacy and behavior as predictors of subsequent behavior in an assertiveness training programme. *Behavior Research and Therapy*, 21, 225-232.

- Lee, C. (1984). Accuracy of efficacy and outcome expectations in predicting performance in a simulated assertiveness task. *Cognitive Therapy and Research*, 8, 37-48.
- Lis, A.; Mazzeschi, C.; Calvo, V.; Salcuni, S.; Parolin, L. (2003) *Psicodiagnostica*, Padova, Unipress (s.d.).
- Lopes R., Bidarra R., (2011) "Adaptivity Challenges in Games and Simulations: A Survey" *Computational Intelligence and AI in Games*, IEEE Transactions on , vol.3, no.2, pp.85-99.
- Luccini, M. A., Mortara M., Catalano C. E., Romero C., (2012) "Thematic Application fields report", Deliverable 3.2 of the Games and Learning (GaLA) Network of Excellence.
- M. Saraò W. L. (2004). *Il modello delle competenze*. Milano: Franco Angeli.
- Marfisi-Schottman, Labat J.-M., Carron T., (2013) Building on the Case Teaching Method to Generate Learning Games Relevant to Numerous Educational Fields. *ICALT 2013*: 156-160.
- Maritain, J., (1947) *La personne et le bien commun*, trans. The Person and the Common Good, by John J. Fitzgerald, Notre Dame, IN:University of Notre Dame Press, 1985.
- Marocco, D., Pacella, D., Dell'Aquila, E., Di Ferdinando, A. (2015) Grounding Serious Game Design on Scientific Findings: The Case of ENACT on Soft Skills Training and Assessment. In Conole, G., Klobu ar, T., Rensing, C., Konert, J., Lavoué, É., *Design for Teaching and Learning in a Networked World*, Lecture Notes in Computer Science, pp 441-446, Vol. 9307. Springer International Publishing, 2105.
- Marocco, D., Pacella, D., Dell'Aquila, E., Di Ferdinando, A. (2015) "Grounding Serious Game Design on Scientific Findings: The Case of ENACT on Soft Skills Training and Assessment", in Conole, G., Klobu ar, T., Rensing, C., Konert, J., Lavoué, É., *Design for Teaching and Learning in a Networked World*, Lecture Notes in Computer Science, pp 441-446, Vol. 9307. Springer International Publishing, 2105. [Link](#)
- Mazzucato, A., Marocco, D., (2015) "Enhancing Negotiation Skills Through On-Line Assessment of Competencies and Interactive Mobile Training", *Open Education Europa*.
- Mazzucato, A., Marocco, D., (2015). *Enhancing Negotiation Skills Through On-Line Assessment of Competencies and Interactive Mobile Training*. *Open Education Europa*.
- Michael, D. & Chen, S., (2006) *Serious games: Games that educate, train, and inform*. Boston, MA.: Thomson Course Technology.
- Mitchell, A., & Savaii-Srnith, C. (2004) *The use of computer and video games for learning. A review of the literature*. The learning and skills development agency. London, p.57.
- Moe, K. O., & Zeiss, A. M. (1982). Measuring self-efficacy expectations for social skills: A methodological inquiry. *Cognitive Therapy and Research*, 6, 191-205.

- O' Broin D., (2011) "Using a Criteria-Based User Model for Facilitating Flow in Serious Games," Third International Conference on Games and Virtual Worlds for Serious Applications, VS-Games.
- Oliveira, M., Cerinsek, G., Duin, H., Taisch, M., (2013) Serious gaming in manufacturing education, In M. Ma et al. (Eds.): Serious Games Development and Applications, SGDA 2013, LNCS 8101, pp. BQ-144, 2013.
- Pacella D., Dell'Aquila E., Yasar Akyar, O., Marocco D. (2014) ENACT: D6.1. Report on artificial cognitive modelling and intelligent tutor: a Literature Review
- Pacella D., Di Ferdinando, A., Dell'Aquila, E., Marocco, D. (2015) "Online Assessment of Negotiation Skills through 3D Role Play Simulation", in Conati, Cristina, et al., eds. Artificial Intelligence in Education: 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings., pp 921-923, Vol. 9112. Springer, 2015. Link
- Pacella D., Di Ferdinando, A., Dell'Aquila, E., Marocco, D. (2015). Online Assessment of Negotiation Skills through 3D Role Play Simulation, in Conati, Cristina, et al., eds. Artificial Intelligence in Education: 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings pp 921-923, Vol. 9112. Springer International Publishing, 2015.
- Parsons, D.; Petrova, K.; Hokyoung Ryu, (2012) "Mobile Gaming -A Serious Business!," Wireless, Mobile and Ubiquitous Technology in Education (WMUTE), 2012 IEEE Seventh International Conference on, vol., no., pp.17,24.
- Pedon A., Gnisci A. (2012). Manuale di psicodiagnostica Aspetti teorici e applicativi dei test. Casa editrice Le lettere – Firenze.
- Pedrabissi, L.; Santinello, M. (1997) I Test Psicologici: teorie e tecniche, Bologna, Il Mulino.
- Pellegrino, J. w., Hilton, M. L. (2012) Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century. Washington, DC: The National Academies Press.
- Poce a., Iovine A., et al., (2014) ENACT: D3.1 Training Needs Analysis Report ENACT Project
- Poce a., Iovine A., et al., (2014) ENACT: D3.2 Research Methodology of the ENACT Project
- Prensky M., (2003) "Digital game-based learning", ACM Computers in Entertainment, vol. 1, no. L.
- Primi C., Chiesi F. (2005). Introduzione alla psicometria. Editori Laterza
- Rahim, A. & Magner, N. (1995). Confirmatory factor analysis of the styles of handling interpersonal conflict: first-order factor model and its invariance across groups. *Journal of Applied Psychology*, 80(1), 122-132.
- Rahim, M. A. (2011). *Managing conflict in organizations*. Third Edition. Transaction Publishers.
- Rahim, M. A., Antonioni, D., & Psenicka, C. (2001). A STRUCTURAL EQUATIONS MODEL OF LEADER POWER, SUBORDINATES' STYLES OF

HANDLING CONFLICT, AND JOB PERFORMANCE. *International Journal of Conflict Management*, 12(3), 191-211.

Rahim, M.A., (1995). *ROCI. Rahim Organizational Conflict Inventory. Manuale. Organizzazioni Speciali, Firenze.* Majer, V.

Riedel, J.; & Baalsrud Hauge, J. (2011) *State of the Art of Serious Gaming for Business and Industry*, 17th International Conference on Concurrent Enterprising: Collaborative Environments for Sustainable Innovation, Aachen, Germany, Centre for Operations Management, RWTH Aachen, Aachen.

Rowe, J. P., Shores, L. R., Mott, B. W., Lester, J. C. (2010) *Individual Differences in Gameplay and Learning: A Narrative-Centered Learning Perspective*. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games (FDG)*, Monterey, CA.

Sawyer B., Smith P., (2008) "Serious Game Taxonomy", Paper presented at the meeting Serious Game Summit 2008, Game Developer Conference.

Schembri M., Marocco D., (2015) *ENACT: D5.1 Report on software development progress and technical specifications Version 2.0*.

Schembri M., Secci A., Di Ferdinando a., Marocco D., (2015) *ENACT: D5.4 software user manual*.

Schwarzer, R., & Jerusalem, M. (1995). *Generalized Self-Efficacy scale*. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, England: NFER-NELSON.

Schwarzer, R., Mueller, J., & Greenglass, E. (1999). *Assessment of perceived general self-efficacy on the Internet: Data collection in cyberspace*. *Anxiety, Stress, and Coping*, 12, 145-161.

Shute, V.J., & F. Ke (2012). *Games, learning, and Assessment*. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in Game-Based Learning: Foundations, Innovations and Perspectives* (pp. 43-58). New York: Springer.

Swarz, J., Ousley, A., Magro, A., Rienzo, M., Burns, D., Lindsey, A.M., Wilburn, B., Bolcar, S., (2010) "CancerSpace: A Simulation-Based Game for Improving Cancer-Screening Rates," *Computer Graphics and Applications, IEEE*, vol.30, no.1, pp.90-94.

Sweetser P. and Wyeth P., (2005) "GameFlow: A Model for Evaluating Player Enjoyment in Games," *ACM Computers in Entertainment*, vol. 3, no. 3.

Sweller, J. (1988) "Cognitive load during problem solving: Effects on learning". *Cognitive Science* 12 (2).

Tukey, J. W (1962) *The Future of Data Analysis*. *Ann. Math. Statist.* 33, no. 1, 1-67

Van de Vliert, E., & Hordijk, J. W. (1989). *A theoretical position of compromising among other styles of conflict management*. *The Journal of social psychology*, 129(5), 681-690.

Watkins, R., Leigh, D., Foshay, R. and Kaufman, R., (1998) "Kirkpatrick Plus: Evaluation and Continuous Improvement with a Community Focus", Educational Technology Research & Development, vol. 46, no.4.

Weider-Hatfield, D. (1988) Assessing the Rahim Organizational Conflict Inventory-II (ROCI-II).

Wendel V., Gutjahr M., Gobel S., and Steinmetz R., (2013) Designing collaborative multiplayer serious games. Education and Information Technologies 18, 2.

Wouters, P., van Oostendorp, H., van Nimwegen, C., & van der Spek, (2013) E. D. A meta-analysis of the cognitive and motivational effects of serious games, Computers and Education, 60, 1.

Zielke M. A., Evans M. J., Dufour F., Christopher T. V., Donahue J. K., Johnson P., Jennings E. B., Friedman B. S., Ounekeo P. L., Flores R., (2009) "Serious Games for Immersive Cultural Training: Creating a Living World," IEEE Computer Graphics and Applications, vol. 29, no.2, pp. 49-60.

Zyda M. (2005) "From visual simulation to virtual reality to games".IEEE Computer.

Zyda, M.; Hiles, J.; Mayberry, A.; Wardynski, C.; Capps, M.; Osborn, B.; Shilling, R.; Robaszewski, M.; Davis, M., (2003)"Entertainment R&D for defense," Computer Graphics and Applications, IEEE, vol.23, no.1, pp.28,36.